







369  
7.9.70











# Psychological Review

THEODORE M. NEWCOMB, Editor  
*University of Michigan*

VOLUME 64, 1957



PUBLISHED BIMONTHLY BY THE  
AMERICAN PSYCHOLOGICAL ASSOCIATION, INC.  
PRINCE AND LEMON STS., LANCASTER, PA.  
AND 1333 SIXTEENTH ST. N. W., WASHINGTON 6, D. C.

Entered as second-class matter July 13, 1897, at the post-office at Lancaster, Pa., under Act of Congress of March 3, 1879

Acceptance for mailing at the special rate of postage provided for in paragraph (d-2), Section 34.40, P. L. & R. of 1948, authorized Jan. 8, 1948

Send all business communications, including address changes, to 1333 Sixteenth St. N.W., Washington 6, D. C. Address changes must arrive by the 25th of the month to take effect the following month. Undelivered copies resulting from address changes will not be replaced; subscribers should notify the post office that they will guarantee second-class forwarding postage. Other claims for undelivered copies must be made within four months of publication.

Copyright © 1957 by the American Psychological Association, Inc.

Harold Edm.      Frederick  
Dated      7. 9. 70  
Copy      8 - 369  
Case No.



## CONTENTS OF VOLUME 64

ALLUISI, E. A. Conditions Affecting the Amount of Information in Absolute Judgments .....	97
ATKINSON, J. W. Motivational Determinants of Risk-Taking Behavior .....	359
BAHRICK, H. P. An Analysis of Stimulus Variables Influencing the Proprioceptive Control of Movements .....	324
BERLYNE, D. E. Uncertainty and Conflict: A Point of Contact Between Information-Theory and Behavior-Theory Concepts .....	329
BROADBENT, D. E. A Mechanical Model for Human Attention and Immediate Memory .....	205
BROWN, W. L. <i>See</i> Overall, J. E.	
BRUNER, J. S. Neural Mechanisms in Perception .....	340
BRUNER, J. S. On Perceptual Readiness .....	123
CARMICHAEL, L. Robert Mearns Yerkes: 1876-1956 .....	1
CLEVELAND, S. E. <i>See</i> Fisher, S.	
DAY, R. H. The Physiological Basis of Form Perception in the Peripheral Retina .....	38
DEMBER, W. N., AND EARL, R. W. Analysis of Exploratory, Manipulatory, and Curiosity Behaviors .....	91
DUFFY, ELIZABETH. The Psychological Significance of the Concept of "Arousal" or "Activation" .....	265
EARL, R. W. <i>See</i> Dember, W. N.	
FELDMAN, R. S. The Role of Primary Drive Reduction in Fixations .....	85
FISHER, S., AND CLEVELAND, S. E. An Approach to Physiological Reactivity in Terms of a Body-Image Schema .....	26
GEORGE, F. H., AND HANDLON, J. H. A Language for Perceptual Analysis .....	14
GIBSON, J. J. Optical Motions and Transformations as Stimuli for Visual Perception .....	288
GUILFORD, J. P. Creative Abilities in the Arts .....	110
HANDLON, J. H. <i>See</i> George, F. H.	
HARLOW, H. F., AND HICKS, L. H. Discrimination Learning Theory: Uniprocess vs. Duoprocess .....	104
HENLE, MARY. Some Problems of Eclecticism .....	296
HICKS, L. H. <i>See</i> Harlow, H. F.	
HOCHBERG, J. E. Effects of the Gestalt Revolution: The Cornell Symposium on Perception .....	73
HURVICH, L. M., AND JAMESON, DOROTHEA. An Opponent-Process Theory of Color Vision .....	384
JAMESON, DOROTHEA. <i>See</i> Hurvich, L. M.	
MALMO, R. B. Anxiety and Behavioral Arousal .....	276
MENKES, A., AND MENKES, J. The Application of Dimensional Analysis to Learning Theory .....	8
MENKES, J. <i>See</i> Menkes, A.	
MILNER, P. M. The Cell Assembly: Mark II .....	242
MOLTZ, H. Latent Extinction and the Fractional Anticipatory Response Mechanism .....	229
OVERALL, J. E., AND BROWN, W. L. Recency, Frequency, and Probability in Response Prediction .....	314
PEREBOOM, A. C. A Note on the Crespi Effect .....	263
POSTMAN, L., AND RILEY, D. A. A Critique of Köhler's Theory of Association ....	61

RESTLE, F. Discrimination of Cues in Mazes: A Resolution of the "Place-vs.-Response" Question .....	217
RESTLE, F. Theory of Selective Learning With Probable Reinforcements .....	182
RILEY, D. A. <i>See</i> Postman, L.	
SIEGEL, S. Level of Aspiration and Decision Making .....	253
SMITH, G. Visual Perception: An Event Over Time .....	306
STEVENS, S. S. On the Psychophysical Law .....	153
TAJFEL, H. Value and the Perceptual Judgment of Magnitude .....	192
UNDERWOOD, B. J. A Graphical Description of Rote Learning .....	119
UNDERWOOD, B. J. Interference and Forgetting .....	49
ZIMMERMAN, D. W. Durable Secondary Reinforcement: Method and Theory .....	373



**Bureau Ednl. Psy. Research**  
**DAVID HARE TRAINING COLLEGE**  
*Dated*.....  
*Acce. No* .....



ROBERT MEARNS YERKES



# THE PSYCHOLOGICAL REVIEW

ROBERT MEARNS YERKES

1876-1956

Robert Mearns Yerkes was born in Breadysville, Bucks County, Pennsylvania, not far from Philadelphia, on May 26, 1876. He died on February 3, 1956, in New Haven, Connecticut. In his death, both psychology and biology lost an outstanding research scientist and organizational leader.

As pictured in his autobiography (20), his childhood was the active one of a boy on a farm. He liked to work with horses and cows. He had many small, wild-animal pets. He collected tortoise and snake eggs, and watched them hatch with keen interest. His uncle, a physician in nearby Collegeville, provided him with work, so that it was possible for him to earn his way through Ursinus College. On graduation, he had intended to study medicine, but an unexpected offer of a loan of \$1,000 for graduate work at Harvard led him to go to Cambridge. At this institution he had the good fortune to study with, and to come to know intimately, most of the philosophers, psychologists, and biologists of the Harvard of that great day. He was a student and special protégé of Josiah Royce. He was also associated with George Herbert Palmer, Francis Peabody, George Santayana, and Ralph Barton Perry among the philosophers. His teachers and associates, among psychologists, included William James, Hugo Münsterberg, Robert MacDougall, and Edwin B. Holt. In the laboratory of zoology, he worked with E. L. Mark, G. H.

Parker, C. B. Davenport, and W. E. Castle. His thesis at Harvard dealt with the sensory reactions and the physiology of the central nervous system of a type of jellyfish.

As student and teacher he stayed at Harvard from 1897 to 1917. He speaks of this period as made up of happy, eventful years of research and teaching. During this time, he created and developed the study of comparative psychology at Harvard, and he also found time for many other important scientific activities. On a brief leave of absence, he acquired knowledge of neurosurgical techniques by working at Johns Hopkins University with the great brain surgeon, Harvey Cushing. Later, he worked with one of his former students, G. V. Hamilton, a medical investigator who, in his private laboratory in California, had the important new idea of working with monkeys to illuminate problems of human behavior and human psychopathology.

Under the inspiration of Dr. Ernest E. Southard, Professor of Neuropathology in the Harvard Medical School, Yerkes worked on a number of problems in human neurology and psychiatry. For five years he gave half his time to the direction of psychological service and research in the Psychopathic Department of the Boston State Hospital, in association with the remarkable group of young students who clustered about Dr. Southard at this time. It was during this period that the Yerkes

Point Scale for Measuring Mental Ability was developed. (14). In 1913, with a former student, D. W. LaRue, he published his *Outline of a Study of the Self* (11). His interest in the relationship of family traits to the personality, as treated in this book, is also attested by the fact that he sent to the National Academy of Sciences, for the file it maintains on each member, a Eugenics Record Office blank completely filled out concerning himself and his near family. In this blank, under the heading "Special tastes, gifts, peculiarities of mind or body, character, favorite pursuits, amusements, etc.," he wrote of himself (in 1912), "Diligent student from youth. Love for research. Left-handed. Poor mechanical memory."

In his fortieth year, in the spring of 1917, he accepted a call to leave Harvard and go to the University of Minnesota to reorganize the psychological work there and take direction of its laboratory of psychology. Among those who were brought to Minnesota by him at this time were Richard M. Elliott, William S. Foster, Mabel Fernald, and Karl S. Lashley. Herbert Woodrow was already at Minnesota at that time. The declaration of war, however, made it seem inappropriate for Yerkes himself to go to Minnesota and, as it turned out, he never physically took up residence there.

In this crisis year he was President of the American Psychological Association. Because of the fortunate fact—for psychology and for America—that he held this office at this time, and because of his energy and his rare gift for the effective and rational planning of organized scientific programs, he saw the challenge that the war could give to psychology and he proceeded to take quick and positive action. More than anyone else, he was responsible for mobilizing almost all of American psychology of the time and placing it at

the service of the nation. With a group of colleagues, he organized and became head of the psychological work and examining in the new and rapidly expanding Army. At first, he was given the rank of major and later promoted to lieutenant colonel. Under his direction, 115 officers and more than 300 trained enlisted men gave intelligence tests to 1,726,966 individuals. Of this number 42,000 were commissioned officers. The work accomplished by this professional psychological group is described in full in the great report, "Psychological Examining in the United States Army," which he edited and which was published by the National Academy of Sciences in 1921 (18).

During this same period Yerkes served as chairman of a committee of psychologists of the newly organized National Research Council. He helped to transform this committee into its present and lastingly effective form as the Division of Anthropology and Psychology of the National Research Council.

Because of his war work, Yerkes saw, possibly more clearly than any other scholar of his generation, the real place and importance of psychology in the service of a complex modern nation. At the end of hostilities, therefore, instead of going to Minnesota, as he was anxious to do, he felt it his duty to stay on in Washington to help in the organization of science in general and of psychology in particular in a nation at peace.

In achieving this objective, he was active in the development of the notable Research Information Service of the National Research Council. This office provided, among other services, the following: a catalogue of the research personnel of the country, a list of research laboratories in industrial establishments, a current list of scientific investigations, an index of published bibliographies of science, a catalogue of major items of scientific apparatus, a list of doctorates



in science in American universities, and many similar services. Owing to a series of misfortunes too complicated to describe here, but related to a postwar national need for financial retrenchment and a failure in certain quarters to achieve a modern vision of the new and truly national place of science in a complex society, this work was eventually terminated. Not until the Second World War, in some functions of the Office of Scientific Research and Development, in the National Roster of Scientific and Specialized Personnel, and later in the activities of the National Science Foundation and the Bio-Sciences Information Exchange of the Smithsonian Institution, were some of these early and constructive administrative ideas again brought into being. The analytic "codes" used by Yerkes in classifying research workers in these early studies were a great step forward in the whole history of the classification of fields of science, as well as of scientific workers.

During his Washington period, under the inspiration of E. W. Scripps of the Scripps-Howard newspapers, and through the instrumentality of Dr. W. E. Ritter, Yerkes participated in the organization of Science Service, the non-profit institution for the proper popularization of science, which still does much to make American newspaper and periodical reporting of science accurate and effective.

During these same active years, he was concerned in the development of two important committees which were organized at the National Research Council. One of these was the Committee on Scientific Problems of Human Migration which was organized under Yerkes' chairmanship in 1922. It grew out of correspondence between Yerkes and certain members of Congress concerning problems of immigration. The committee considered, from the point of

view of natural science, the complex migrational situation resulting from the World War. It also prepared and undertook the coordination of research programs intended to provide reliable information on the physical, mental, and social characteristics of ethnic groups. Another committee was that for Research in Problems of Sex. At the first meeting of this committee, in 1922, Yerkes was elected chairman. He held this post for 25 years. During this time the committee expended \$1,456,745, largely in grants for scientific research. It is recorded that 585 individuals participated in this work.<sup>1</sup>

In 1924, after this amazingly active and fruitful period in Washington, and seven years after leaving Harvard, he accepted a professorship in the Institute of Psychology at Yale, the predecessor of the Institute of Human Relations at that University. The great development of these institutes was encouraged by the distinguished psychologist, James R. Angell, who was then president of Yale.

A study of Yerkes' publications, from his first paper on the light reaction of certain crustaceans in 1899 (1), shows in a most dramatic way the depth and also the amazing breadth of his scientific and administrative accomplishments.<sup>2</sup> Many of these papers deal with the sensory and neural processes of a wide variety of living organisms. Special attention is given to the speed of reaction, to inborn behavior, and to the learning process in these publications. Among the most important of these early contributions are those

<sup>1</sup> S. D. Aberle, & G. W. Corner. *Twenty-five years of sex research: history of the National Research Council Committee for Research in Problems of Sex, 1922-1947*. Philadelphia: Saunders, 1953.

<sup>2</sup> *Psychological Register*, Vol. 3. Worcester: Clark Univer. Press, 1932; *Publications from the Yerkes Laboratories of Primate Biology, Inc.*, Orange Park, Fla. (rev. to 1953).



which deal with habit formation in the crab (2), the sense of hearing in the frog (3), and objective nomenclature in comparative psychology and animal behavior (4). A book, *The Dancing Mouse* (5), and a paper with S. Morgulis on the method of Pavlov (6), were possibly especially significant. This latter paper, published in 1909, was most important in introducing the idea of the conditioned reflex to English-speaking readers.

In 1911, he published, with John B. Watson, the often-quoted monograph on methods of studying vision in animals (7). In the following year, his well-known paper on the intelligence of earthworms appeared. In 1913 his work on the heredity of savageness and wildness in rats did much to start a whole tradition of research in psychology (12). About this time, also, he published a paper on color vision in birds (15).

Interspersed with these technical and psychobiological studies are papers in what may be called quantitative applied psychology. This illustrates a fact not always recognized in psychology, that the thoroughly trained comparative psychologist has many techniques at his command which are especially useful in psychotechnology, or what is now called human engineering. In Yerkes' case this tendency is illustrated by his study of the psychological aspects of illuminating engineering (8), and in his experiments on the psychology of advertising materials (10).

In his Harvard period he wrote on the relation of psychology to medicine (13), and indeed much of his time there was given to this topic, as illustrated in his development of new approaches to the measurement of mental capacity, and in his studies of mental examinations of police court cases (16). He published a textbook, *Introduction to Psychology* (9), in 1911. In 1916 his notable

monograph on the mental life of monkeys and apes was issued (17).

After he went to Yale in 1924, his papers began to show a special preoccupation with the broad and significant problems of psychology, biology, and medicine, which he saw as related to the life of the chimpanzee and the other great apes. Possibly the most notable publication of his life is *The Great Apes: A Study of Anthropoid Life* (19). In reference to this work he writes in his autobiography: "In 1905, when I was fairly started in my career as a psychobiologist, I began a partnership with Ada Watterson (Yerkes), which perfectly blended our lives and incalculably increased our professional and social usefulness. Successful marriages appear in these times to be not unworthy of record and remark. Moreover, from 1905 my professional autobiography is no longer mine alone. At this moment our partnership is publishing jointly, as the outcome of six years of continuous preparatory labor, a book on anthropoid life, *The Great Apes*" (20, p. 391).

In the months between the close of his administrative work at the National Research Council and his assumption of duties at Yale, he devoted the summer of 1924 to research in anthropoid behavior in Havana in the primate colony established there by Mrs. Rosalia Abreu.

Once located in New Haven, Yerkes threw himself with energy into the establishment of a special laboratory for the full psychobiological study of the great apes. The whole growth of this idea is set forth in his book, *Chimpanzees: A Laboratory Colony* (21), published in 1943, which describes the history of the scientific study of the chimpanzee and also the development of the laboratories which were essentially Yerkes' own creations. In the epilogue of this volume, he notes that in 1900, while he was still a graduate

student at Harvard, the idea of a special research institute for comparative psychobiology suddenly came to him. From that time on, in spite of war and administrative duties, he worked earnestly to bring into being something which, in its completeness, was altogether new in the world. This novel conception was an adequate research institute for the comparative study of behavior, with special emphasis on the investigation of problems related to the great apes.

In 1925, with the cooperative support of Yale and the Rockefeller Foundation, he was able to begin to test out, in fact, the feasibility of keeping and using chimpanzees in New Haven for experimental purposes. At length, further funds became available for the realization of his dream of the establishment of a laboratory in a region more favorable for the life of these great animal subjects than was New England. After detailed study, a subtropical site at Orange Park, Florida, was purchased. This location, about 15 miles from the city of Jacksonville and near the St. Johns River, has proved to be most fortunate. In 1930 the erection of basic buildings began here, and the nucleus of the chimpanzee colony was established. When Dr. Yerkes retired as Director of these laboratories in 1941, it could be said with assurance that he had demonstrated unequivocally the great scientific importance of this unique primate laboratory in its subtropical climate. Yerkes described the success of this work under the following five headings (21): (a) The laboratory had shown itself to be capable of successful breeding and rearing of experimental animals of known ancestry and history. (b) Practical problems related to feeding, housing, and hygiene involving these large, intelligent, but very destructive animals, had been worked out. (c) Much basic information con-

cerning the anatomy, physiology, and psychology of the individuals of the colony had been acquired. (d) A system of laboratory and colony records had been established which provided a biography of each animal. (e) Special studies had been carried out dealing with growth, maturation, sexual cycle, auditory, visual, and perceptual processes, known correlates of behavior, discriminative learning, development of the use of tools; study of symbolism, ideation, and insight; a study of linguistic expression and capacity, of emotional traits, social behavior, drug addiction, and many other problems.

After Dr. Yerkes' retirement in 1942, the name of the Orange Park laboratories was appropriately changed, in his honor, from the Yale Laboratories of Primate Biology to the Yerkes Laboratories of Primate Biology, Inc.

This is not the place to evaluate in detail the many important contributions made at the Yerkes Laboratories. A list of the workers who have been trained at this station, and a study of upwards of 400 research papers that have been issued there, attest the significance that this institution has had, not only in furthering modern scientific study of the higher apes, but also in the study of the neural and general physiological basis of behavior. At the time of his death Dr. Yerkes was still an active member of the Board of Scientific Directors of the Laboratories. Those who have served with him on this Board can all attest to the wise counsel that he had given to the Laboratories since his retirement as Director. This active support meant much to Dr. Karl S. Lashley, who succeeded him as Director. At Lashley's retirement in 1954, Dr. Yerkes was pleased by the selection of Dr. Henry W. Nissen as Director. It is a cause for real satisfaction that he lived to see the distinguished work done at the Laboratories under Lashley's



direction and the fine scientific programs now being conducted there.

During the Second World War, Yerkes was active in consulting services in the War Department and at the National Research Council. For him, the second war brought again into a new focus some of the old problems with which he had wrestled so successfully during the First World War. One of the committees of the Division of Anthropology and Psychology of the National Research Council during the Second World War was the Emergency Committee in Psychology. Dr. Yerkes served as chairman of one of its most important and fruitful subcommittees—that on survey and planning. Out of the work of this group grew the notable Constitutional Convention at which the present divisional structure of the American Psychological Association was formulated. Yerkes served as temporary chairman of the convention. The constructive work of this meeting in establishing an organization which held American psychology together as a single body, instead of allowing it to dismember itself into separate organizations, was in no small measure the result of his wisdom and long-tested administrative skill.

Dr. Yerkes was elected to membership in the National Academy of Sciences in 1923, and to the American Philosophical Society in 1936. As already noted, he was president of the American Psychological Association in 1916–17. He was president of the American Society of Naturalists in 1938. Among other organizations in which he was active are the American Academy of Arts and Sciences, the American Physiological Society, and the Society of Mammalogists. He was a member of the Cosmos Club of Washington, D. C. He received honorary degrees from Ursinus, Wesleyan, and Yale, and in 1954 he received the New York Zoological Society's gold medal for his pioneer work in psychobiology.

He is survived by his widow, the former Ada Watterson; a son, David Norton Yerkes of Washington, D. C.; a daughter, Roberta Watterson Yerkes of New Haven; and a brother, William Augustus Yerkes of Southampton, Pennsylvania.

Dr. Yerkes never retired. Up to the very time of his death, he was actively concerned in thinking about problems of science in relation to comparative psychology and the improvement and survival of *Homo sapiens*. In 1945 he spoke on chimpanzees as servants of science, on one of the notable nationwide series of radio talks of the New York Philharmonic program (22). In 1946 he published a thoughtful, forward-looking paper on psychology in world reconstruction (23), and in 1951 a "gorilla census" (24) which focused attention on the need for more study of that primate. His last publication is entitled "The Biologist's Point of View" (25). Certainly the title of this paper may almost be taken as the motto of his life and work. In everything that he did in the study of lower organisms or of the higher apes, as well as in his dealing with human beings in his many and important administrative associations, he always thought and acted as a humane natural scientist dealing with complex independent units of organized protoplasm.

One sentence from this last paper well summarizes the guiding faith of this distinguished scientist and great man: "Even though relatively few of us may be scientists occupationally, everyone can and should be scientific in spirit and understanding" (25).

#### REFERENCES

1. Reaction of Entomostraca to stimulation by light. *Amer. J. Physiol.*, 1899, 3, 157–182.
2. Habit formation in the green crab, *Carcinus granulatus*. *Biol. Bull. Woods Hole*, 1902, 3, 241–244.



3. The sense of hearing in frogs. *J. comp. Neurol.*, 1905, 15, 279-304.
4. Objective nomenclature, comparative psychology and animal behavior (editorial). *J. comp. Neurol.*, 1906, 16, 380-389.
5. *The dancing mouse: a study in animal behavior*. New York: Macmillan, 1907. Pp. xxi + 290.
6. (With S. Morgulis). The method of Pavlov in animal psychology. *Psychol. Bull.*, 1909, 6, 257-273.
7. (With J. B. Watson). Methods of studying vision in animals. *Behav. Monogr.*, 1911, 1, No. 2. Pp. vi + 90.
8. The psychological aspects of illuminating engineering. *Lectures on illuminating engineering* (delivered at Johns Hopkins University, Oct.-Nov., 1910, Vol. 2). Baltimore: Johns Hopkins Press, 1911. Pp. 575-604.
9. *Introduction to psychology*. New York: Henry Holt, 1911. Pp. xii + 427.
10. The class experiment in psychology with advertisements as materials. *J. educ. Psychol.*, 1912, 3, 1-17.
11. *Outline of a study of the self*. Cambridge: Harvard Univer. Press, 1913. P. 24.
12. The heredity of savageness and wildness in rats. *J. anim. Behav.*, 1913, 3, 286-296.
13. Comparative psychology in relation to medicine. *Boston med. surg. J.*, 1913, 169, 779-781.
14. *A point scale for measuring mental ability*. Baltimore: Warwick & York, 1915.
15. Color vision in the ring-dove (*Turtur risorius*). *Proc. nat. Acad. Sci., Wash.*, 1915, 1, 117-119.
16. Mental examination of police and court cases. *J. crim. Law Criminol.*, 1916, 7, 366-372.
17. The mental life of monkeys and apes: a study of ideational behavior. *Behav. Monogr.*, 1916, 3, No. 1. Pp. iv + 145.
18. Psychological examining in the United States Army. *Mem. nat. Acad. Sci.*, 1921, 15, vi + 890.
19. (With Ada W. Yerkes). *The great apes: a study of anthropoid life*. New Haven: Yale Univer. Press, 1929. Pp. xix + 652.
20. Robert Mearns Yerkes, *psychobiologist*. In *A history of psychology in autobiography*, Vol. II. Worcester: Clark Univer. Press, 1932. Pp. 381-407.
21. *Chimpanzees: a laboratory colony*. New Haven: Yale Univer. Press, 1943. Pp. xv + 321.
22. Chimpanzees as servants of science. (One of a series of radio talks, N. Y. Philharmonic program, June, 1945.)
23. Psychology in world reconstruction. *J. consult. Psychol.*, 1946, 10, 1-7.
24. Gorilla census and study. *J. Mammalogy*, 1951, 32, 429-436.
25. The biologist's point of view. *Animal kingdom* (N. Y. Zoological Soc.), 1954, 57, No. 2, 34-35.

LEONARD CARMICHAEL

Smithsonian Institution

## THE APPLICATION OF DIMENSIONAL ANALYSIS TO LEARNING THEORY

AVIVA MENKES AND JOSH MENKES<sup>1</sup>

*University of Michigan*

Dimensional analysis is an analytical method by means of which one can obtain functional relations between variables in a physical problem. The variables must be capable of being reduced to an expression in terms of a fundamental set (in mechanics, mass, length, and time). This paper has the twofold purpose of (a) proposing a fundamental set for some psychological variables, and (b) applying the method of dimensional analysis to these variables.

This mode of analysis rests on the premise that an equation expressing the relationship between a number of variables has meaning only if these variables have the same dimension, where a dimension is any product of the fundamental set.<sup>2</sup> (In mechanics this would be  $M^a L^b T^c$ , where  $a$ ,  $b$ , and  $c$  are positive or negative integers or zero.) In simple terms, this statement asserts that one cannot add apples and bananas unless one "reduces" both to fruit. An equation in which the variables have the same dimension is called *dimensionally homogeneous*. As this concept is fundamental to dimensional analysis, it will be expounded in some detail.<sup>3</sup>

The term homogeneous as applied to algebraic equations implies that each term in the equation has a com-

mon factor, say,  $x$ . This factor can be divided out and the equation will still hold true, e.g.,  $5x - 4xy = 0$ . This equation is said to be homogeneous in  $x$ .<sup>4</sup> Equations involving dimensions, where a dimension is the qualitative characteristic of an object, are no exception. (Object is used here as a member of a class containing at least two terms which can be characterized both quantitatively and qualitatively.) Consequently,  $x$  could be a dimension and then the equation would be said to be homogeneous in the dimension  $x$ , or simply dimensionally homogeneous. If  $y$  were considered to be a dimension, the equation would of course be meaningless. In a dimensionally homogeneous equation we can cancel the dimensions and are left with a relation that involves only pure numbers, and is thus independent of the method of measurement that was employed in establishing the equation in the first place. Since the laws of arithmetic apply only to relations of pure numbers, it is clear that in order to manipulate our equations we have to make sure that we handle equations that are dimensionally homogeneous. This requirement is the foundation on which we will build our analysis.

It is now necessary to demonstrate that dimensional homogeneity is a requirement to be fulfilled by all functions containing psychological variables.

<sup>1</sup> In the Department of Aeronautical Engineering.

<sup>2</sup> This is a purely mathematical definition of dimension. An equivalent definition is: A dimension is the qualitative aspect of a quantity.

<sup>3</sup> A detailed account of the method and its uses can be found in Bridgman (3) and Langhaar (5), both of which require a minimal amount of mathematical background.

<sup>4</sup> There exist more rigorous definitions of homogeneity for equations of higher degree, which, however, are unnecessary for this paper.

The relationships between two psychological variables can be essentially of two natures. It can be (a) in the form of an equation expressing the functional relation between variables, or (b) of a correlational type. In the first case, one can equate two classes of objects only if both exist and both contain *equivalent elements* (6, pp. 183-185). As an object is characterized by a quantity and a dimension, an equation must by definition be dimensionally homogeneous. In the second case, it could be argued that one could fit a curve to experimental points and consider the curve as describing the relationship of the two variables, maintaining that nothing in the operations performed actually demanded dimensional homogeneity. Consider the simple case where the relation was found to be approximately linear. The curve fitted to such points would be of the form:  $y = a + bx$ , and let us assume furthermore that  $y$  and  $x$  have different dimensions but have been demonstrated to be related to each other. However,  $b$  is nothing but  $dy/dx$  and  $a = y_0$ , i.e., the value of  $y$  at  $x = 0$ . Rewriting the equation as  $y = y_0 + (dy/dx)x$ , we immediately recognize it as a dimensionally homogeneous one. In case the relationship is not linear, but is "well-behaved"<sup>5</sup> (a most plausible assumption to make) the function can be expressed as a Taylor Series, which, by the very mode of its construction, is dimensionally homogeneous.<sup>6</sup>

<sup>5</sup> A "well-behaved" function is one which changes smoothly and continuously. This does not exclude oscillatory and similar functions.

<sup>6</sup> Any "well-behaved" function of an arbitrary number of variables can be expressed as a Taylor Series. As a matter of fact, the mathematical definition of well-behavedness is equivalent to the statement that the functions can be expanded as a Taylor Series. Every standard treatise on the calculus can be consulted on details about the Taylor Series.

One is thus forced to conclude that the functions of psychology should also fulfill the requirements of dimensional homogeneity. There are two implicit assumptions involved in this presentation. The first is that the variables are measurable by at least an interval scale. (The scaling method that dimensional analysis generates will be shown later.)<sup>7</sup> The second assumption is that deterministic laws can be found in the particular area of investigation. The possibility of describing the phenomena probabilistically is not rejected. As to which of the two approaches is the more fruitful one, only time will tell, and in the meantime we find comfort in a statement by Einstein in a letter to Max Born. "You believe in the dice-playing God, and I in the perfect rule of law. . . . I am absolutely convinced that one will eventually arrive at a theory in which the objects connected are not probabilities . . ." (2).

The necessity for dimensional homogeneity imposes certain restrictions on the relationship in which variables can appear in an equation. Thus by knowing the variables but nothing about their relationship to each other, one can by some very simple manipulations find their functional relationship.<sup>8</sup> Upon considering the applicability of this method to psychology, one is confronted by the problem that no independent members of a fundamental set, or basic dimensions, have been ascribed to psychological variables. The place of dimension in

<sup>7</sup> Coombs has developed the notion of four types of models—the compensatory, conjunctive, disjunctive, and lexicographic. Dimensional analysis is a method usable only on compensatory models. Thus, if the method does not turn out to be useful, the possibility that the variables under consideration do not act in a compensatory manner is always possible.

<sup>8</sup> See footnote 3.



physics can be illustrated by the following quotation:

In the domain of mechanics we are called upon to deal with various aspects of nature to which we give such names as length, mass, time, force, velocity, momentum, energy, etc. *By common consent* (our italics) we take the first three of these as basic and independent. . . . If we then fix units for the measurement of these three basic quantities, we are immediately in a position to express the measure of any other quantity . . . in terms of these basic units (5).

We now have to proceed by analogy from the physical sciences and assume the existence of a fundamental set of variables (which are measurable *directly* or *indirectly* with a certain degree of precision) such that all other variables in the particular area of investigation can be expressed in terms of this fundamental set. A purely hypothetical example will illustrate this point; if tension and attention are considered basic dimensions, then perceptual vigilance or defense could be stated as changes in attention with respect to tension.

Dimensional analysis can thus be used for the following purposes:

1. Given the variables that are relevant in a particular situation, and given the dimensions of these variables, a function can be found which expresses the relations of all these variables to each other *without* experimentation, only by means of some simple analysis.

2. In case the variables known to be important in a particular situation cannot be combined into a dimensionally homogeneous equation, it can be construed as a sign that an important variable is missing—or certain variables could be neglected. As the dimensions of the supposedly missing variables may suggest themselves upon examination, this has an important heuristic function. A simple example from mechanics will be introduced so that we will not have to deal in ab-

stract generalities. It was known from experimental investigations that

$$R \propto S/a,$$

$R$  being the resistance caused by two bodies sliding over each other, say submersed in oil, which is proportional to the relative speed  $S$ , and the distance of separation  $a$ . The dimensions of these variables are  $R = F/L^2$ ,  $S = L/T$  and  $a = L$ . This equation can also be written as

$$R = K \times S/a,$$

where  $K$  is called the constant of proportionality, whose dimension must be chosen so as to make the equation dimensionally homogeneous. It will obviously have to be  $TF/L^2$ . This variable was called the coefficient of viscosity, and was found later to be an important variable in physical problems. In addition, the dimensions of this variable can be fully justified on the basis of fundamental laws of mechanics and make good *physical* sense.

3. Dimensional analysis can be used to scale psychological variables. The discussion of this application will be postponed to a later point in the paper until a specific example from the psychology of learning can be utilized.

Dimensional analysis can *not* tell the investigator which variables are functioning in a situation (except in the sense of point 2, above). It is *not* a theory—in the sense of a guide for the selection and definition of variables. It is merely an analytical tool, useful once the selection of variables and the assigning of dimensions to these variables has been accomplished.

The important thing to remember is that there is absolutely nothing about dimensions that is God-given or intuitively apparent, apart from the fact that dimensions are variables in terms of which all other variables can be expressed.

### Application to Hull's Variables

Now that the method has been explicated the afore-mentioned purpose of the paper will be attempted. It is fully realized that the dimensions arrived at are only tentative, and might turn out to be quite fruitless, though this would not necessarily imply the lack of utility of the approach.

The variables in the following equation will be expressed in terms of *drive* and *time*, which will be considered as basic dimensions.

$$s\bar{E}_R = f(sH_R, V, J, K, D, sI_R, I_R) \quad (4).$$

Reinforcement itself is not included as a variable but, as it plays an important role in the other variables, dimensions will be assigned to it. As drive reduction occurs in an interval of time, reinforcement can be thought of as a change of drive with respect to time, i.e., it is expressible as a derivative  $\frac{dD}{dT}$  or  $DT^{-1}$ .

$K$  is the incentive, and its amount is assumed to determine the slope of the reinforcement curve; it can therefore be expressed as the second derivative  $\frac{d^2D}{dT^2}$  or  $DT^{-2}$ .

$V$  represents the stimuli evoking the response. It was found very difficult to define their dimension in terms of their function, but as Hull assumes them to attain secondary reinforcing value, they will, provisionally and with misgivings, be considered as having the same dimension as reinforcement.<sup>9</sup>

$sH_R$  is a function of the number of reinforcements, and its dimension is therefore the same as that of reinforcement itself.

$J$ , the delay, has an effect that is opposite to  $K$ . It can also be thought of as determining the slope of the reinforcement curve, and its dimension is therefore the same as that of  $K$ .

<sup>9</sup> See footnote 3.

$I_R$  and  $sI_R$  can be thought of as making for an increase in drive, which lessens the strength of  $sE_R$ . Their dimension is therefore equivalent to reinforcement, though opposite in its effect.

$sE_R$  will be considered to have the dimension of time because it is measurable directly or indirectly in terms of latency scores.

To summarize the dimensions of the variables:

$$\begin{aligned} sE_R &= T & K &= DT^{-2} \\ sH_R &= DT^{-1} & J &= DT^{-2} \\ sI_R, I_R &= DT^{-1} & D &= D \end{aligned}$$

These variables may now be grouped into nondimensional clusters, a nondimensional cluster being a group of variables which is a pure number, e.g.,  $\frac{sH_R}{sI_R}$  or dimensionally  $\frac{DT^{-1}}{DT^{-1}}$ .

Some of these groupings might turn out to be important in experimental situations, i.e., the variables function interdependently—a change in one with a concomitant change in the other (so that the value of the nondimensional cluster remains constant) should result in no change of an independent variable.

Two such nondimensional ratios may be plotted against each other by systematically varying only one vari-

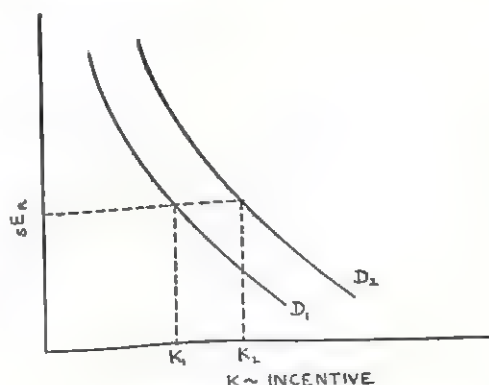


FIG. 1. Illustration of scaling technique.

able in one and obtaining measures for only one variable in the other, and thus achieving a two-dimensional graphical representation of more than two variables.

Now the use of dimensional analysis in scaling can be shown. If a cluster has been demonstrated to be truly nondimensional, then it can be plotted as shown in Fig. 1. If  $D_1$  is six hours of deprivation and  $D_2$  is twelve hours of deprivation, and we assume the one to be twice the amount of the other—i.e., we use these as a base line—then  $K_1$  can be considered to be twice the amount of  $K_2$ . Or, more cautiously, the relationship of the magnitude of  $K_1$  to  $K_2$  is the same as that of  $D_1$  to  $D_2$ .

With the above-mentioned variables the following nondimensional clusters were formed:

$$\frac{KD}{J^2 \times (sE_R)^2}; \frac{sH_R}{sI_R}; \frac{K}{J}; \frac{sH_R \times I_R}{V \times sI_R}.$$

Whether these are truly nondimensional will have to be determined experimentally.

Assuming these clusters to be nondimensional, the following combination into an equation is possible:

$$\frac{sE_R^2 \cdot K}{D} = C_\alpha \left( \frac{K \cdot I_R \cdot sH_R}{J \cdot sI_R \cdot V} \right)^\beta$$

$$sE_R = \sqrt{\frac{D}{K}} \left( C_\alpha \left[ \frac{K \cdot I_R \cdot sH_R}{J \cdot sI_R \cdot V} \right] \right)^\beta.$$

The value of the constant and the exponent will have to be determined experimentally. This is not the only possible equation. Other dimensional homogeneous equations are equally acceptable; however, the number possible is highly restricted. At this point such predictions are obviously premature, but quite feasible after some systematic investigation.

### *Preliminary Experimental Investigation<sup>10</sup>*

Some evidence that the present analysis might be adequate was gained by the following pilot study.

The purpose of the experiment was to establish whether the variables  $K$  and  $D$  were compensatory as the above analysis would suggest. (See the nondimensional cluster

$$\frac{K \times D}{J^2 [\text{Const.} = sE_R^2]}.)$$

*Procedure.* It was first necessary to establish whether  $D$  and  $K$  (of which preference was considered to be the index) were independent of each other. Three different saccharine solutions were presented to two groups of albino rats ( $N = 24$ ) under conditions of 24 and 8 hours of water deprivation. They had free access to all three solutions, and the amount consumed of each was recorded. Both groups preferred the same solution, namely, that found optimally preferred by Beebe-Center (1). It was therefore concluded that  $D$  and  $K$  were independent.

In the experiment proper, two groups of albino rats were trained in a Skinner box. For each bar-pressing a drop of fluid was released into a plate. The animals were trained until their average pressing rate was constant over a 3-day period. One group was trained under 6-hour water deprivation and presented with the optimally preferred saccharine solution, the other being trained under 22 hours' deprivation and reinforced with plain water.

The  $sE_R$  measure taken was rate of response after it had leveled off. The mean response for the two groups was

<sup>10</sup> The authors would like to express their sincere thanks to Dr. Russell Clark who guided the design and execution of these tests.



TABLE 1  
RATE OF BAR-PRESSING IN SKINNER BOX  
UNDER TWO EXPERIMENTAL CONDITIONS

	Group I	Group II
	8 hrs. deprivation saccharine solution reinforcement	16 hrs. deprivation water reinforcement
Mean	4.08	4.06
SD	.45	.38

calculated, the results being as shown in Table I.<sup>11</sup>

*Conclusion.* The difference between the two groups would occur by chance 100% of the time as  $t = 0$ . It was therefore considered reasonable to accept the null hypothesis.

#### SUMMARY

The method of dimensional analysis was presented. Its uses as a method of combining variables into functional

<sup>11</sup> A more detailed report of experimental procedure and findings can be obtained upon request.

relationships by analytical considerations, as a way of discovering new variables, and as a tool for scaling were discussed. Some variables of the psychology of learning were dimensionalized and analyzed by means of the above methods. These considerations yielded a number of predictions, one of which was subjected to a preliminary test. The results were as predicted.

#### REFERENCES

1. BEEBE-CENTER, J. G., BLACK, P., HOFFMAN, A. C., & WADE, M. Relative per diem consumption as a measure of preference in the rat. *J. comp. physiol. Psychol.*, 1948, 41, 239-251.
2. BORN, M. *Natural philosophy of cause and chance*. Oxford: Clarendon Press, 1948.
3. BRIDGMAN, P. W. *Dimensional analysis*. New Haven: Yale Univer. Press, 1922.
4. HULL, C. L. *Essentials of behavior*. New Haven: Yale Univer. Press, 1951.
5. LANGHAAR, H. L. *Dimensional analysis and theory of models*. New York: Wiley, 1951.
6. STEBBING, L. S. *A modern introduction to logic*. London: Methuen, 1930.

(Received September 12, 1955)

## A LANGUAGE FOR PERCEPTUAL ANALYSIS

FRANK H. GEORGE

University of Bristol, England

AND

JOSEPH H. HANDLON<sup>1</sup>

College of Medicine, University of Illinois

The aim of this paper is to develop further the theory of behavior that has already been outlined by the authors in a previous publication (3).<sup>2</sup> In particular, it aims to develop that part of the total theory which deals with *perception* and to clarify what has sometimes been called the "nativist-empiricist controversy." The method we have adopted in presenting the over-all theory is to proceed by degrees toward a state of greater precision by selecting for elaboration at each stage some single aspect of the whole theory. It is at present a skeleton theory, where we are attempting to lay down a *language* to deal with the principal problems of behavior. The aim is as much meta-theoretical as theoretical, and involves the clarification of existing terminology as well as the laying down, albeit tentatively, of rules for the *use* of such terminology.

### THE MODEL: GENERAL BEHAVIORAL UNIT

The behavioral model centers around the notion of a *belief*. It is suggested that a belief is ultimately to be investigated in terms of the states of the central nervous system; but it may be thought of, in the Peircian manner, as "that which disposes the organism to

act" (8). Belief is here used as a theoretical term, a sort of logical construct.

Beliefs are considered to be relatively permanent states of the central nervous system. In everyday terms, they are those stored memories whose contents specify for the organism what may be expected to happen ( $S_2$ ) if certain behavior ( $R_1$ ) is performed under certain stimulating circumstances ( $S_1$ ). Since at any given moment the organism's behavior is a function of a relatively few of the totality of its stored beliefs, we shall call those beliefs which are actually influencing behavior at any given instant of time *expectancies* ( $E$  ( $R_1 - S_2$ ))s, or more briefly,  $E$ s). Beliefs may be converted into expectancies through the action of the *activating stimulus* state ( $S_1$ ). This activating stimulus state is a conjunction of *motivational stimuli* ( $S$ ms), stemming from the motivational state ( $M$ ) of the organism, and of the *stimulus complexes* ( $S^*$ s).

Both the  $S$ ms and the  $S^*$ s are subclasses of the class of stimuli that have been associated with particular beliefs.<sup>3</sup>  $S$ ms are, of course, internal to the organism, while the  $S^*$ s may be either external or internal. One possible subclass of  $S^*$ s (the relation may actually be one of overlap, class exclusion, or

<sup>1</sup> United States Public Health Service Post-Doctoral Fellow in Clinical Psychology.

<sup>2</sup> The authors wish to thank Dr. John W. Yolton of the Department of Philosophy, Princeton University, for his continued helpful advice, especially with this paper.

<sup>3</sup> We have previously defined "stimulus" as "any change of energy which impinges upon the nervous system such that it brings about activity of that system. The source of this energy change may be either external or internal to the organism" (3, p. 31). Obviously, not all such stimuli have been associated with particular beliefs.

class inclusion) is the class of *modifying motivational stimuli* (MMs) which are capable of changing the internal *motivational state* ( $M$ ). This motivational state, which is seen as being composed of two factors, *drive* ( $D$ ) and *urgency* ( $U$ ), may act to determine the range of the expectancies transformed from beliefs.<sup>4</sup>

After a range of expectancies has been transformed from the totality of stored beliefs by the activating stimulus state ( $S_1$ ), a hypothetical process is postulated by which the range of expectancies is *scanned*. This process of scanning leads to the "selection" of a single expectancy whose correlative response ( $R_1$ ) is the one that will be subsequently emitted. The "selection" of the single expectancy during the scanning process is made in terms of (a) *the strength of the belief* underlying the expectancy, and (b) *the valence* of the expectancy. Valence, in turn, is a function of the *anticipated reward* and the *anticipated effort* involved in the projected response. The emission of the correlative response ( $R_1$ ) associated with the chosen expectancy follows automatically upon the selection of that expectancy. This response will either be followed by the *anticipated outcome* ( $S_2$ ), in which case either a *confirmation* of the correlative belief will take place, or it will not, in which case a *falsification* of the belief will follow. This whole process, beginning with the transformation of beliefs into expectancies by the  $S_1$  and ending with confirmation or falsification, is called the *general behavioral unit*.

While the interactions of the above

<sup>4</sup> "Motivation" has been defined by us as "a state of the organism, compounded of drive and urgency factors, which, through the medium of motivational stimuli ( $Sms$ ), produces behavior in conjunction with stimulus complexes ( $S*s$ ) through the elicitation, rehearsal, and selection of expectancies; and the elicitation of responses" (3, p. 33).

variables are too complex to be presented in any adequate detail here, a few further points should be made. Which beliefs are converted into expectancies depends upon the previous association between certain stimuli and certain beliefs. Beliefs are thus acquired by contiguity—the association in experience of activating stimuli ( $S_1s$ ) with means outcomes ( $R_1 - S_2s$ ). Such experiences may be direct ones, wherein physical energy changes, forming the basis for the stimuli, emanate from that event about which the organism acquires beliefs. But many beliefs are learned indirectly through the use of symbols, where the emanating energy changes from the event do not in any direct fashion determine the beliefs about that event. Thus, knowledge may be acquired either through *direct acquaintance* or through *description*.<sup>5</sup> Apart from their acquisition, beliefs may be *strengthened* (through confirmation) or *weakened* (through falsification), either by description or by direct acquaintance, by going through the steps of the general behavioral unit (GBU).

Motivation is not only an important factor in the determination of the range of expectancies elicited for scanning, but it also plays an important role in the selection process through its indirect effect upon the valence. Further, motivation operates to determine the speed of selection as well as the speed and strength of response elicitation.

There are other behavioral units called the *perceptual behavioral units* (PBUs) which are closely related to GBUs, and are necessary to them. For every GBU there must exist a PBU to categorize the originally registered *activating stimulus state* ( $S_1$ ). A PBU is also necessary to identify the *outcome* or *goal* ( $S_2$ ) to allow assessment of *confirmation* or *falsification*. It is clear

<sup>5</sup> It has been suggested that certain stimulus-belief associations may be innate as well.



that a PBU must precede every GBU and must also follow every GBU. The possible exception is in so-called "thinking," where GBUs might follow each other.

#### THE MODEL: PERCEPTUAL BEHAVIORAL UNIT

We shall now turn our attention to the *perceptual behavioral unit* (PBU) itself and note, first of all, the similarities and differences between it and the previously described *general behavioral unit* (GBU).

In general, the basic form of the PBU is similar to that of the GBU. After proper encoding, certain of the stimuli impinging on the central nervous system are capable of transforming beliefs into expectancies; then, through the process of scanning, one of the expectancies is selected, and a response ensues. But there are certain important differences between the details of the GBU and the PBU that must be noted.

1. The first difference concerns itself with the nature of the final response ( $R_1$ ). In the PBU the  $R_1$  is the covert response of categorizing or classifying the impinging stimuli. Such a response is best thought of as occurring entirely within the central nervous system, and as not necessarily involving conscious awareness. In contrast, the  $R_1$  of the GBU may be any number of different actions, some of them overt (motor activity) and some covert (emotional changes, "thinking").

2. We must distinguish between the *contents* of the beliefs of the PBU and those of the GBU. In the PBU, beliefs concern themselves with such cognitive actions as seeing, hearing, tasting, etc.—in general, those activities which have traditionally come under the rubric "sensation and perception." They involve the action of the central nervous system as the organism "apprehends" its external, as well as its internal, en-

vironment. Perceptual beliefs can be expressed as conditionals of the form: if the impinging stimuli (the  $S_1$ ) have been categorized as  $C_1$  (the  $R_1$ ), then the subsequent impingement by other stimuli of the categories  $C_2 \dots C_n$  (the  $S_2$ ) is likely to obtain, with the probability  $P$ . What the conditions are under which these probable impingements will take place also forms the content of perceptual beliefs—or, more precisely, the perceptual meta-beliefs.

For convenience's sake, we shall consider all those beliefs that are concerned with the perceptual categorization of events to be *perceptual* beliefs, and all others *general* beliefs. That such an arbitrary division as this is only a temporary verbal convenience will be appreciated when we now note the close relationship between GBUs and PBUs.

3. In the GBU, in order for an activating stimulus state ( $S_1$ ) to transform a belief into an expectancy, it is first necessary that the  $S_1$  be *perceived*. By perception, of course, we mean the action of a categorizing response which is the  $R_1$  of an immediately preceding PBU. That is to say, in order for a GBU to take place it must be preceded by a PBU. On the other hand, in the PBU, in order for an  $S_1$  to transform a perceptual belief into a perceptual expectancy we must arbitrarily assert that it is *not* possible that the  $S_1$  be preceded by a categorizing response. (Indeed, if in the present theory this *were* permitted, we should obviously be involved in an infinite regress.) Thus, for the PBU, given a particular  $S_1$ , those perceptual beliefs that are associated with the  $S_1$  will immediately and automatically be converted into perceptual expectancies. This transformation will take place without the  $S_1$ s being first perceptually categorized. The only categorization involved is the  $R_1$  which ends the PBU.

4. Since there may be more than one

perceptual belief associated with a given  $S_1$ , the question arises as to *which* of the beliefs converted into perceptual expectancies will be selected. Such a selection will lead automatically to the categorizing response,  $R_1$ , of the PBU. We have seen that in the GBU the selection of the expectancy which leads automatically to the  $R_1$  is a joint function of (a) the valence of the expectancies which have been transformed from beliefs, and (b) the strength of the beliefs correlative with these expectancies. In the PBU, however, while these two factors of valence and belief strength also operate to select a single expectancy, it is the latter that seems to play the more important role. This does not mean that valence may not be important, especially when the stimulating circumstance ( $S_1$ ) is ambiguous and when motivation is strong; but, on the whole, the strength of the perceptual belief is primary in the selection of a particular perceptual expectancy leading to the  $R_1$  of the PBU. By strength of perceptual belief we mean, simply, the degree or strength of association existing between a particular  $S_1$  and a particular perceptual belief.

5. Unlike the GBU, the final step in the PBU is not an *outcome* or *goal* ( $S_2$ ) which is then followed by another PBU leading to a categorizing response, for this would now involve us in an infinite progress. Rather, the PBU ends with an  $R_1$ , the categorizing response. This perceptual categorization may, however, be subsequently confirmed or disconfirmed; and this, in turn, will lead to the strengthening or weakening of the correlative perceptual belief through confirmation or falsification. Such confirmation or falsification may take place in two ways. First, the outcome ( $S_2$ ) of the response ( $R_1$ ) of the subsequent GBU may either confirm or falsify the veridicality of the previous perceptual categorizing response. Secondly, a sub-

sequent PBU, because of the content of the organism's belief system, may be categorized as being either compatible or incompatible with the previous PBU in question. And this, in turn, may bring about a confirmation or falsification of the previous PBU, leading to the strengthening or weakening of the correlative perceptual belief. It should be noticed that perception is normally regarded as *certain* by the organism, and is not thought of as requiring further confirmation.

We must distinguish between beliefs regarding perceptual *events* and beliefs regarding perceptual *rules*. We might also talk of perceptual beliefs regarding *rules about rules*. Another way of distinguishing these various levels of perceptual beliefs is to call them "beliefs," "meta-beliefs," and "meta-meta beliefs," etc.

Let us now continue our examination of the PBU in detail. We have previously defined the activating stimulus state ( $S_1$ ) as "... that state of the central nervous system which is capable of transforming specific beliefs into expectancies" (3, p. 32) and we have considered the  $S_1$  to be made up of stimulus complexes ( $S^*$ s) plus stimuli arising from motivational states ( $Sms$ ). But in order to specify more precisely the functioning of the  $S_1$  in the PBU, it is convenient to partition  $S_1$ s into three subcategories: (a) cues, (b) clues, and (c) signs.

We shall postulate, as part of the connotation of the words "cue," "clue," and "sign," that they refer to the organism's *use* of certain stimuli *after* they have been associated with specific perceptual beliefs. Thus, we shall consider cues, clues, and signs to be a subclass of the class of activating stimulus states ( $S_1$ s) rather than, say, a subclass of stimulus complexes ( $S^*$ s) or motivational stimuli ( $Sms$ ). For we wish to make it clear that cues, clues, and signs



have their functional basis not only in events external to the organism, but also in events internal to it, such as beliefs, attitudes, motivations, etc.

1. CUES (*Cus*): *A subclass of the class of activating stimulus states ( $S_1s$ ) which stems from objects or events, either internal or external, which are being apprehended directly by the organism through knowledge by acquaintance.*<sup>6</sup>

Cues from external events may be modified by internal *Sms* stemming from concomitant motivational states. Conversely, cues from internal events may be modified by *S's* stemming from concomitant external events.

2. CLUES (*Cls*): *A subclass of activating stimulus states ( $S_1s$ ) which stems from the context, ground, or surround in which the apprehended object or event, either external or internal, is imbedded. The apprehension of a clue by the organism is direct, i.e., through knowledge by acquaintance.*

Unlike cues, clues are not apprehended in and of themselves; if they are experienced at all, they are experienced in conjunction with cues, and they usually exert influence upon the apprehension of the concomitant cue, and vice versa.

Clues, then, are  $S_1s$  which inform the organism about some events or objects other than themselves. Such information may have to do with such "objective" matters as the size, color, shape, location in space, etc., of objects as well as such "subjective" matters as pleasantness, attractiveness, harmfulness, etc. Such information as is given by clues may or may not, of course, be veridical.

3. SIGNS (*Sns*): *A subclass of the*

*class of activating stimulus states ( $S_1s$ ) which the organism, through the acquisition of beliefs, has learned stand for other stimuli.*

As can be seen by the definition, signs are closely related to clues, since they are a subclass of  $S_1s$  which are concerned with objects or events other than themselves. But, unlike clues, which must always appear with the object or event about which they yield information, signs may give information about objects or events which are not simultaneously present. In this sense, clues may be considered to be a kind of subclass of the class of signs.

As we see from the above definition, signs may come to stand for, or be substitutable for, or come to represent objects or events to the organism which are not now present. In fact, under special circumstances, they may come to stand for events or objects which can never be known directly (i.e., through knowledge by acquaintance) by the organism. Thus, though the information the clue conveys is always about an object or event known through knowledge by acquaintance, in the case of signs this need not be so. This is particularly true of signs that are encountered in language (7) which are used in formal and informal education to convey information about objects or events which may never be directly experienced by the organism.<sup>7</sup> Cues as well as clues may, of course, act as signs under special circumstances.

This is by no means a sufficient analysis of signs, since that part of the glossary definition which says they "stand for" itself requires a great deal of further explanation. There is no space for this analysis here.

<sup>6</sup> See the authors' previous discussion (3) for the distinction between knowledge by description and knowledge by acquaintance. "Apprehended" is a neutral term that can be easily defined as a function of the  $S_1$ . That is, it is simply the registering of a stimulus by the central nervous system—not necessarily in consciousness. As used in this paper, "experience" has a similar meaning.

<sup>7</sup> While it is theoretically possible for all but the very first sign-object association to be learned through knowledge by description, in most actual cases the child learns most of his early language signs by experiencing directly both the sign and the object, usually contiguously in time.



Objects and events about which cues and clues convey information can be experienced only directly, while the objects or events about which signs convey information can be experienced either directly or indirectly. Knowledge *about* cues, clues, and signs, however, may be acquired either through knowledge by acquaintance or knowledge by description.

The breaking down of  $S_1$ s into the above subclasses may often turn out to be rather arbitrary; nevertheless, it is believed to be a convenient way of describing the various functions of the activating stimulus state in the PBU. The point of breaking down the integrated act of *perception*, in the way described, is to avoid the impasse of classical Gestaltism which allows too little of the necessary analytical processes of science. At the same time our reconstruction is intended to cater to the wholeness principle of Gestaltism which has been too often overlooked by classical behaviorism.

Often, when we apprehend an object or event, we are not able to categorize it completely or fully, to our satisfaction, at the first attempt. Instead, we may first categorize it one way and then another, and so on. We may, of course, take into account each one of our "abortive" categorizations to form a concatenation of interpretations from which we construct our final categorizing response. Or we may make use of further incoming information from the object or event itself. Or, finally, by examining our past experiences, through our perceptual beliefs, we may try to "remember" what this object or event "might be." We then arrive at a categorization upon which we are prepared to act, insofar as we are relatively sure that our categorization is a "correct" one.<sup>8</sup>

<sup>8</sup> The word "categorization" is intended by us to mean both (a) the activity of separating any two parts of "sensory time-space," and

From this we see that we can look upon the perceptual process as consisting of a finite series of interpretations:  $I_1, I_2, \dots, I_n$ , where  $I_n$  is that interpretation or categorization upon which the organism is prepared to act. In the limiting case, the series of interpretations may, of course, be only two— $I_1, I_2$ , or even merely  $I_1$ . The series of categorizations may take place extremely rapidly, and it is only on relatively rare occasions that such a series is slow enough and perhaps difficult enough, so that we become conscious of this process.<sup>9</sup>

It is therefore necessary, for any complete description of the perceptual process (or process leading to perception), to make allowances for such a series of interpretations or categorizations. We shall, therefore, distinguish between what we will term the *provisional categorizing responses* (PCRs) and the *final categorizing response* (FCR), the latter being of particular interest since it forms the  $S_1$  of the subsequent GBU.

From what we know of visual perception, those categorizations and interpretations of "the given" ( $S$ 's and  $S$ ms) which occur early in the total percep-

---

(b) the subsequent activity of drawing up classifications in terms of similarities and differences. In this sense, the word has a broader meaning than is usually intended.

<sup>9</sup> The reader should not infer from what has been said that we are here advocating any return to the structuralist position in which perception is compounded of elementaristic units bound together by the "cement" of association. But the rejection of such a position is not to deny the piecemeal, serial nature of so many of our perceptions. Further, when operating within a linguistic model, such as the one we are proposing, such an analytic tool as a series of terms, not necessarily functionally independent, seems absolutely necessary. Thus, there is an important relationship between the analytic (logical) and the psychological meaning of "interpretation." Any such discussion of this relationship, such as would be elaborated in pragmatics, would not be appropriate to this paper.

tual process are primarily concerned with such matters as the separation into figure-ground, form, color, texture, etc., leading to categorizations of basic recognition. Whether these categorizations and interpretations which occur early in the perceptual process are acquired through experience, as suggested by Von Senden and Hebb (10, 4), or whether they are part of the built-in functions with which we are born (6), is an issue which cannot be settled at this time. We are here merely pointing out the fact that certain perceptual processes are prior in time to others. The later categorizations and interpretations, besides correcting earlier misperceptions, tend to be concerned with such factors as the perceptual context, elaborated meanings and implications, etc. Here the earlier categorizations are combined, elaborated, and given added connotations from past experience through the action of the organism's belief system. We assume, of course, that the same process is duplicated in all the sense modalities. That such elaborations may be enormously complex can be attested to by the difficulty encountered when one attempts to specify the exact sequence of events of an aesthetic experience involved, say, in the reading of a poem or the contemplation of a painting.

As hinted at above, often it is the categorization of the context in which the object or event is imbedded that may greatly influence the FCR of that object. For example, suppose we place a straight stick in water such that half of it is submerged beneath the surface. Now, if we look at the stick from a certain angle, we may very well categorize it as being "bent" if we have not yet observed the stick to be half submerged in water, or if we are not aware of the effect that media of differing refractive indices have upon the path of light. Once, then, we are aware of the context, and have appropriate beliefs about the

influence of the context upon the object, we may categorize the stick as being "straight" even though we might continue to say that it "looks bent."

We must, therefore, at times distinguish between the way a particular object or event is *sensed* and what the object or event is *believed* to be. In terms of the present model, the earlier *sensing* of the object would correspond to the interpretations of the earlier of the PCRs, while the final *believing* of what the object or event is would correspond to the FCR. The stick in water may be *sensed* as looking bent, but it will be *believed* to be a straight stick half immersed in water.

Both *sensing* and *believing* involve the action of beliefs, expectancies, and the selection of a categorizing response, so that these two processes are really not different in kind, but are simply a matter of relative position in a series of interpretations.

We may, of course, cite other perceptual illustrations of the bent-stick type in which the PCRs are not always in agreement with the FCR. These illustrations are often used to make the point that experience may have very little effect upon what is perceived (9). Our disagreements with such an interpretation will be elaborated upon in the next section; nevertheless, these illustrations are helpful in demonstrating certain important details of the proposed model. For example, looked at from the top of the Empire State Building, cars on the street below may be said to "look like" (be *sensed* as) toy autos, but they will be "believed to be" normal-sized cars observed from a great height. Or, standing on the ties of a railroad right of way, the rails may "look like" (be *sensed* as if) they converge at the horizon, but are "believed to be" parallel. In these rather special cases, stimulus configurations which we have previously learned to categorize in



one particular way may not correspond with the actual situation in the present instance. Indeed, an important part of one's "maturing perceptually" is the learning of those contexts where it is the better part of wisdom to disregard or make exceptions to previously valid perceptual beliefs and meta-beliefs. On the other hand, *most* perceptual situations are such that the perceptual beliefs we hold do have generalizability, insofar as they have a validity which is trans-situational, and in which the PCRs *do* correlate highly with the FCR. In this sense, then, the examples cited above are, despite their popularity in psychological theorizing, not representative of usual perceptual situations (1).

#### THE SEMANTIC PROBLEM OF PERCEPTION

The analysis of the last section has some immediate applications with respect to the controversy between those who contend that experience has little or no influence upon what is perceived (the nativists) and those who contend that experience is a major determiner in what is perceived (the empiricists). This controversy can now be revealed for what it primarily is—a pseudo-problem. It is a pseudo-problem in the sense that it is not empirical; that is to say, the problem has been created by the use of certain terms in various different and ambiguous ways. Both empiricism and nativism reflect some aspects of the truth, and their differences have hinged on two different interpretations of the same set of terms. These different interpretations have led to the examination of highly selected data. Within this selection, confirmation has been found for some aspects of the respective theories. The arguments have thus taken on a certain circularity, and the result has been an increasing terminological confusion.

We have already stated the essentials

of perception for our system, and the vitally important fact that arises is that the processes which have been variously and vaguely labeled *sensing*, *seeing* (or *hearing*, *touching*, etc.), *perceiving*, *believing*, and so on, refer to the complex stages of the operation which starts (say) with the impinging of a visual stimulus on the retina, and might conventionally be said to be completed by some action leading to confirmation or falsification on the part of the organism. In our terms, it begins with *S*'s and *MMS*s, and ends with either an overt or covert response. The confusion has arisen in the past over just how these processes should be described in detail. Indeed, this is our problem: to construct a suitable terminology, or linguistic framework, for the proper discussion of these perceptual problems. This is obviously not the whole answer to the problem of perception, but it will help to avoid confusion over the description of perception and this is virtually the whole answer to most nativist-empiricist disputes.

Now we should further consider some examples that were discussed in the last section. Let us consider again a stick half immersed in water. If the observer is asked how the stick is *perceived*, he may be at a loss as to how to answer correctly. This is, of course, because the word "perceive" is vague, and may be taken to mean any subset of the set of PCRs, even including the FCR. Thus if he says he perceives the object as a bent stick, it implies that the *cues* have been adequately categorized, but not all the *clues*, and the word "perceive" is taken to apply to some subset of PCRs that have not sufficiently fulfilled their full clue function, in that they fail to tell the observer that, since the stick is immersed in water, in reality the stick is probably straight.<sup>10</sup> By

<sup>10</sup> It should be noticed that the clues, cues, and signs present have not fulfilled their full





"in reality," we mean in the ordinary sense of being seen under some standard conditions and with respect to (say) tactual clues. If, then, the observer took a slightly broader interpretation of "perceive," he might have included such a clue about the immersion in water, and then have said that he perceived a straight stick partially immersed in water. Of course the answer to the original question is a terminological matter, and depends on the manner in which the word "perceive" is interpreted. Exactly the same difficulties apply to the other cognitive words such as "sensing," "seeing," and "believing."

It is difficult to avoid the conclusion that nativists and empiricists have disagreed mainly over how the word "perception" should be used, and have accordingly carried out experiments which support their own particular usage. Thus, for example, Gestaltists have tended to choose abstract stimulus complexes which are almost wholly devoid of experiential factors. They have thus limited the number of clues to the minimum, while the empiricists have generally done the opposite.

Let us now consider some further examples—first, the distorted rooms of the Ames demonstrations (5). If, under the specified conditions, an observer looks at a distorted room and reports it as being normal, he may be making this statement with respect to the full range of PCRs (right up to and including the FCR); i.e., he really *believes* it is normal. He has, in the words of C. S. Peirce, "no reason to doubt" that it is normal. It may be argued that this is because he lacks the cues and clues

that tell him of the distortion. This, then, is *not* an analogous case to the stick immersed in water, where his *beliefs* may more obviously be different from his *sensings*, i.e., the clues *are* observable in the stick-in-water problem, but *not* in the case of the distorted rooms. But then it is quite apparent to any rational observer that a specially distorted room (distorted, that is, with an eye to deceive) could look normal, if viewed from a particular angle; so the reason why his interpretation (or belief) is that the room is normal is that normal-appearing rooms, in experience, *are* normal. Thus he believes this both as a result of his sensings and as a result of the rest of his relevant experience.

Now we are faced with a more difficult question. Does the observer actually *see* events differently, as a result of the experience, or does he merely *believe* these events to be different? Once again we are in terminological difficulties over how broadly or how narrowly to interpret words such as "see" in terms of PCRs. In natural language the word is vague, and there is no reason to suppose that any observer can wholly divorce what he *sees* from what he *believes*. Indeed, such evidence as we have (for example, from trained artists) strongly suggests that this is a relationship which can be greatly changed with careful training.

Thus we seem to be forced to the conclusion that there is no sensible test that can be applied to the situation where the experimenter says, "Report only what you see." This is the case both because of the vagueness over words like "see," and because of the nonverbal difficulties observers have in separating one subset of the PCRs from the total set. The most obvious answer to this psychological dilemma is to concentrate on an investigation of the sensory systems from a strictly physiological point of view. This indeed is the

---

function relative to some categorizing decision if the FCR is incorrectly made. That is, cues, clues, and signs have "intension" (sense) and "extension" (denotation), and all that is implied may not be extracted by the observer, although complete extraction *may* not be necessary to a correct decision.

culminating point of this analysis. It points up the near impossibility of progressing further in perceptual experiments without recourse to physiological hypotheses, and it also points out the serious difficulties attendant upon the verbal instructions in any perceptual experiment.

There are other examples, which have already been discussed in Section II, that have been quoted by writers of nativist inclinations to show the lack of need for experiential factors in perception. Let us consider the cases of "parallel railroad tracks" and "cars viewed from the top of the Empire State Building." These examples bring out the difficulties well. In each case a process of sensing (primitive recognition) takes place, and two converging lines and some tiny cars may be said to be *seen*. With respect to the parallel lines there is a clear distinction—as with the stick in water—between what is *seen* and what is *believed*. Nobody would expect (in the ordinary sense of the word) to see railway tracks as parallel from the back of a train, if by the phrase "see railway tracks as parallel" we mean "see two lines that appear to stay equidistant from each other throughout their length." One of the *meanings* (i.e., one of the strong beliefs that exist about them) of parallel lines is that they should appear to converge from certain viewpoints. Similarly with cars; one expects them to look tiny from the top of the Empire State Building. And incidently, the philosopher's popular example is also analogous: you expect a circular coin to look roughly elliptical from certain angles.

We are certainly aware that these problems cause linguistic and logical difficulties for the philosophical theories of perception; but what problem do they offer for psychology? There is no problem. No one claims that we should see such railroad tracks as parallel, i.e., as

*appearing* equidistant, as part of them would from perpendicularly above.<sup>11</sup> Nor indeed has anyone argued (to the best of our knowledge) that experience would make you *see* (in this sense) such things. The whole argument is over the word "see." You may *see* (in one literal usage of the word obviously near the sensing end of the continuum) the lines as converging, and yet you may also *see* (in another sense near the believing end of the continuum) the tracks as parallel. It is absolutely clear here that the word "see" is being taken in two different senses. In this verbal confusion both nativists and empiricists must accept some guilt.

The processes we are concerned with are the series processes (the set of interpretations  $I_1, I_2, \dots, I_n$  referred to above), which we may say start with *sensing* and run through to *believing* (the successive PCRs). It is certainly clear that, in all the examples quoted, the beliefs depend, as it were, on the presence of, and the making of the correct inductions with respect to sufficient cues and clues to interpret the object and the context of the object viewed. It cannot, however, be said, even in *apparently* purely visual problems, that such beliefs necessarily depend on either visual or nonvisual clues exclusively. Nor is it always possible to separate out the *sensing* from the *believing*. For example, under artificial circumstances, such as in the use of some abstract designs used in the laboratory, or by the use of some reduction-screens in the case of the constancies, our clues are seriously curtailed, and we are forced to *believe* what we literally *sense*. Under natural circumstances where we view

<sup>11</sup> The reader should be very careful with the word "see" as we have used it here. We mean that the retinal image should not be projected onto the visual cortex as lines equidistant from each other at all points. Such a state is not possible if the lines are anywhere parallel in fact.



cars from the top of the Empire State Building, railroad tracks from the back of trains, sticks partially immersed in water, and so on, we have clues in abundance which allow us to interpret in a manner which is obviously a function of experience. The point here is, of course, that by "interpretation" we finally mean "with respect to a belief" and beliefs are built up in experience; thus in our model, the function of experience, with respect to clues, is axiomatic. However, this should not be taken to obscure the fact that experiential modifications may also be built into the sensory processes themselves, nor that certain perceptual beliefs may be innate.

The best that can be suggested at the moment is that such terms as "sensing," "seeing," "perceiving," and "believing" be used with constant awareness of their great vagueness. We shall use the terms PCR and FCR as part of what we define as the PBU, and shall generally try to avoid the use of the above terms except with special care. This implies a careful statement as to the range of *meaning* that such terms carry with respect to any particular context. For purposes of constructing scientific models, we still lack "natural" (physiologically given) criteria for the partition of the subsets of the total process. This is a specific and important problem that requires solution. In the absence of any differentiating criteria, we can only point to a particular conventional usage that we intend at any given time. Thus the process beginning with stimulation, continuing through the transformation of beliefs into expectancies, and ending with categorization, is all taken by us to be part of the process that we shall call perception. We shall thus suppose that, in our sense of the term "perception," it is indeed by very definition dependent on experience, as

well as on the limitations and vagaries of the structure of the sensory process.<sup>12</sup>

There is one important point that should be made clear with respect to the nativist-empiricist controversy. In distinguishing "seeing" from "sensing," to take one example, we recognize that we *can* define these words as *meaning* anything we please, but in trying to explicate them we have assumed a difference which sometimes appears to exist in ordinary usage and then suggested how such a difference could be given a sensible explication. The nativists and empiricists in their quarrel may ascribe any *meanings* they choose to their cognitive terms, but they *must* formulate their disagreements over empirical facts in the *same linguistic framework*. This we claim they have not done, nor even recognized that they have not done so. Thus if they could have agreed on, or even recognized, external questions (about their framework), they would have found there were no internal questions (2) of the kind that have appeared in their discussions.

#### REFERENCES

1. BRUNSWIK, E. *Systematic and representative design of psychological experiments*. Berkeley: Univer. of California Press, 1947.
2. CARNAP, R. Empiricism, semantics and ontology. In Linsky, L. (Ed.), *Semantics and the philosophy of language*. Urbana: Univer. of Illinois Press, 1952.
3. GEORGE, F. H., & HANDLON, J. H. Toward a general theory of behavior. *Methodos*, 1955, 7, 25-44.
4. HEBB, D. O. *The organization of behavior*. New York: Wiley, 1949.

<sup>12</sup> This section has so far contained no reference to the word "sign." The discussion has been concerned with learning by acquaintance, and it will be appreciated that signs themselves are directly experienced. Indeed, clues and cues may themselves be signs; and also, signs which are neither cues nor clues occur in perceptions. This notwithstanding, it is in learning by description that signs (as the term is used by us) most frequently occur.



5. KILPATRICK, F. P. (Ed.) *Human behavior from the transactional point of view*. Hanover, N. H.: Institute for Associated Research, 1953.
6. KOFFKA, K. *Principles of Gestalt psychology*. New York: Harcourt, Brace, 1935.
7. MORRIS, C. W. *Signs, language, and behavior*. New York: Prentice-Hall, 1946.
8. PEIRCE, C. S. *Collected papers of . . .* C. Hartshorne and P. Weiss (Eds.). Cambridge: Harvard Univer. Press, 1931-1935.
9. PRATT, C. C. The role of past experience in visual perception. *J. Psychol.*, 1950, 30, 85-107.
10. VON SENDEN, M. *Raum- und Gestaltfassung bei operierten Blindgeborenen vor und nach der Operation*. Leipzig: Barth, 1932.

(Received November 22, 1955)

## AN APPROACH TO PHYSIOLOGICAL REACTIVITY IN TERMS OF A BODY-IMAGE SCHEMA<sup>1</sup>

SEYMOUR FISHER, AND SIDNEY E. CLEVELAND

*VA Hospital, Houston, Texas*

### PROBLEMS IN CATEGORIZING PHYSIO- LOGICAL MEASURES

It is now possible to measure with considerable precision a large number of physiological reactions in the body. One may obtain an almost endless variety of indices of physiological reaction, such as blood pressure, heart rate, finger temperature, saliva output, GSR, and respiration rate. A problem that has confronted those (1, 4, 22, 28) who are interested in such indices is to find a schema which will permit a meaningful grouping or clustering of the indices. That is to say, if one obtains ten different measures of physiological reaction from each of a group of subjects, can one functionally equate certain of the measures and not simply treat them as if they were ten disparate variables? There has been a tendency in the past to lump together the separate measures and to regard them grossly as an overall index of emotionality (3, 4). This was based on the implicit assumption that emotional response is a diffuse, unpatterned process, just a general stirring up of many physiological systems. This point of view was concerned mainly with what happens to the organism after it has been disturbed and its usual equilibrium upset. It has had relatively little to say about the pattern of physiological measures obtained from the undisturbed organism. It also has had little to say about the relationship of physiological

measures to such psychological constructs as "personality" and "traits."

A more differentiated frame of reference which has been much used in analyzing physiological reactions has revolved about the difference between sympathetic and parasympathetic response. There are obvious anatomical and chemical reactivity differences within the autonomic system that lend themselves to the sympathetic-parasympathetic distinction. In an overly simplified fashion one may conceptualize the sympathetic subsystem as functioning mainly with reference to threats to the individual, whereas the parasympathetic subsystem appears to function mainly to conserve and maintain supplies for the individual's needs and development. After Gaskell (cited in Little [25]) published his views regarding antagonistic innervation (sympathetic vs. parasympathetic) of most internal organs, there developed an increasing interest in the use of the concept to explain a wide variety of phenomena. Attempts were made to classify various illnesses as being primarily due to sympathetic or parasympathetic dominance. That is to say, it was assumed that an ideal balance between the two should exist, and that if one became dominant over the other, this imbalance would lead to organ changes which would be pathological (12). Attempts were also made to apply this imbalance concept to persons without clinical symptoms: it was assumed that one or the other subsystem might be somewhat dominant in normal individuals, and cause them to develop certain personality tendencies (36). In addition,

<sup>1</sup> We wish to acknowledge special gratitude to Dr. Hebbel E. Hoff, Chairman, Department of Physiology, Baylor Medical College, for valuable criticisms and suggestions concerning a number of aspects of this paper.

there were efforts to take all known measures of physiological response and classify them as being indices primarily of sympathetic or of parasympathetic activity. Generally, there was little objective basis for most of the assumptions that were made regarding the relationship of the autonomic subsystems to either symptom patterns or normal personality characteristics. There was even some vagueness about what kinds of physiological measures were specifically representative of one autonomic subsystem as contrasted to the other.

More recently, Wenger (40) and others (22, 25) have approached the problem of autonomic functioning with more quantitative and controlled methods than those previously used. Wenger indicates, on the basis of his analysis of the interrelations among many physiological variables in children and in adults, that an autonomic factor score can be extracted which expresses the degree of imbalance between sympathetic and parasympathetic functioning. He has found some evidence that the direction and degree of imbalance may be correlated with degree of personality disturbance and also with a number of normal personality variables. But the degree of relationship is low and variable. It is also noteworthy that he did not find it possible to differentiate patients with stomach ulcers from normals, on the basis of autonomic imbalance scores. This is an important point, because there have been many attempts in the past (16, 35) to describe this and other psychosomatic syndromes in terms of sympathetic-parasympathetic imbalance. Wenger's inability to distinguish ulcer patients from normals in terms of autonomic imbalance has been confirmed by Little (25). One of the most interesting aspects of Wenger's work is that he has shown not only the limitations of the sympathetic-parasympathetic concept, but has also discovered

other factors which are important in accounting for the variance found in any array of physiological measures. In his factor analysis of a wide range of physiological indices obtained from normals under stress and nonstress conditions, Wenger isolated not only an autonomic factor but also other moderately well-defined factors. His results indicate that the problem of categorizing physiological reactivity is a very complicated one. Darrow (9, 10) and Cattell (4) have also found that arrays of physiological measures cannot meaningfully be accounted for in terms of one or two simple dimensions.

Malmo and Shagass (27, 28), Lacey (22, 23), Lewinsohn (24), and Wolff (43) have further complicated the problem of categorizing measures of physiological reactivity by their demonstration that subjects may display fairly consistent patterns of physiological reactivity which are unique to the individual. Their work has indicated that one individual may respond to a stress with maximal activation of a given physiological function, moderate activation of another function, and underactivation of still another function. A second individual, however, may show a hierarchy of activation which is quite the reverse in character. These patterns of response cannot be explained suitably in terms of gross concepts like "sympathetic" or "parasympathetic." Those who have noted these patterns have tried to associate them with personality constructs. It has been especially theorized that some of these patterns are correlated with the same constellations of forces that result in an individual's developing a psychosomatic symptom in one organ system rather than in another (27, 28); it has been hypothesized that just as certain kinds of psychosomatic symptoms may be one of the products of a given style of personality defense, just so may an individual's pattern of physi-



ological reactivity be significantly affected by his style of life. In both instances it is assumed that life experiences have fixated certain modes of body response, and the psychosomatic symptom is simply a more extreme example of such fixation. Several investigators (23, 25, 43) have indeed demonstrated experimentally that there is a significant tendency for individuals to manifest their maximum physiological reactivity to a variety of stimuli in those body systems that also are the focus of psychosomatic symptomatology. This whole line of thought has brought to the fore the importance of a psychological frame of reference in viewing physiological phenomena.

#### BODY-IMAGE CONCEPTS

The problem arises, then, as to what kinds of psychological schemas we can develop which will aid in understanding an array of physiological measures.

The schema to be suggested in this paper has grown out of a series of studies of body image that we have completed (6, 7, 13, 14). We have shown that individuals differ considerably in the attitudes they take toward their own bodies. More specifically, we have found that individuals vary in the degree to which they conceive of their body exteriors as having an armoring protective function. Some visualize their body boundaries as thick, impermeable defensive barriers. Others visualize their boundaries as thin and easily penetrated. These body-image fantasies were measured by means of two Rorschach indices which have been shown to be reliably scorable and also valid in terms of a whole series of criteria (13, 14). One of these indices, the "Barrier" score, is the sum of all responses in which there is an emphasis on the thickness, strength, or unusual covering quality of the periphery. It is illustrated by such percepts as "cave with rocky walls," "mummy wrapped up," and "man in

armor." The second index, referred to as the "Penetration of Boundary" score, is the sum of all responses in which there is an emphasis upon the periphery as penetrated, broken, or fragile. The following are examples of such responses: "bullet penetrating flesh," "butterfly with tattered wings," "plucked chicken." In order to eliminate the effect of response total upon the two body-image scores, it is requested of subjects that they give a fixed number of responses to each blot. Details of this technique are available elsewhere (13).

Our work has demonstrated that a number of personality variables are significantly linked with the degree to which the individual emphasizes the barrier or armored quality of his body boundaries. Thus, our data (14) indicate that persons who visualize their boundaries as thick and armored manifest a higher level of aspiration, more drive toward self-expression, and more motivation for competitive advancement of the self than do persons with indefinite body-image boundaries. At still another level, we have shown that patients hospitalized with psychosomatic symptoms involving the body exterior emphasize the barrier quality of their body boundaries more often than do patients with psychomatic symptoms involving the body interior. For example, it was shown that patients with symptoms like neurodermatitis, which involves the skin, and rheumatoid arthritis, which involves the joints and striate muscle sheath of the body, greatly emphasize the barrier aspects of their body-image boundaries. However, patients with interior symptoms, like stomach ulcers or ulcerative colitis, are distinguished by the permeability that they attribute to their body boundaries. By means of suitable control groups, it was shown that such differences in body image were not merely the result of experiencing symptoms in contrasting

areas of the body. Thus, a group of individuals with skin defects due to industrial accidents was compared with the neurodermatitis group; and a group of patients with low back pain due to accidental injuries was compared with the arthritics. If the mere experiencing of symptoms in the exterior layers of the body determined the individual's body-image fantasies, one would expect the two groups with symptoms due to accidents to have as high Barrier scores as the two psychosomatic groups. Actually, it was found that each of the psychosomatic groups significantly exceeded its respective control group in Barrier responses.

Since Lacey (23), Malmo and Shagass (27), and others (43) have suggested that the same factors may be operating in the development of a psychosomatic symptom in a given organ system as are involved in the manifestation of a given hierarchy of physiological reactivity, the possibility presented itself of applying the body-interior vs. body-exterior schema to the existing literature concerning physiological reactivity. That is to say, since this body-image schema had proven meaningful relative to a variety of psychosomatic symptom phenomena, it seemed logical to examine its potentialities relative to problems of physiological reactivity. The schema is not intended to stand in opposition to other modes of classifying physiological response, but is offered for consideration as an additional dimension from which to view such response. In its application to physiological reactivity the schema assumes that although a given physiological response may be triggered by certain centers (e.g., the hypothalamus), various sectors of the body may react differentially to the excitation.<sup>2</sup> Specifically,

it is being assumed that the individual's life experiences may be one factor which results in his responding to excitation differentially so far as his body exterior and body interior are concerned. Body exterior is here defined as including the skin and striate musculature. Body interior is considered to include all of the internal viscera. This definition of body interior vs. body exterior is intended to have purely location or geographic implications, and is not at all concerned with the embryonic origins of various body areas.

The idea that differences in physiological reactivity may be linked not with conventionally defined organ systems or nervous system divisions, but rather with local characteristics of different body areas has derived support from an elaborate study carried out by Davis, Buchwald, and Frankmann (11). They investigated a whole gamut of physiological reactions in subjects under a variety of stimulus conditions, and attempted to define the patterns that stood out. The patterns of physiological response they found did not correspond to any of the usual modes of categorization in the literature. Indeed, they coined new terms to designate the observed patterns. In describing the concept they have evolved of what constitutes a pattern or system of physiological response to a stimulus they say:

... it is evident that a system in this sense does not correspond with any organ system of the body nor yet with the gross structural divisions of the nervous system. The receptors, the neural structures, and the effectors involved are each activated by a prior element in the chain, and, except for the ultimate one, activate the next succeeding element. *Local characteristics*<sup>3</sup> of the elements will determine the character of the action, transducing it from one sort of force to another. . . . One is tempted to compare the spreading to the propagation of waves in a homogeneous medium—

ferences in the patterning of central response which result in differential stimulation of various body areas.

<sup>3</sup> Italics are ours.

<sup>2</sup> It is also possible to conceive of differential exterior and interior reactivity as a function of selective central processes. That is to say, it could be assumed that there are individual dif-



a better picture would be that of a branching network, which the nervous system actually is, with the possibility of transmission at any point being able to vary somewhat independently of that at any other.

The body-image schema proposes that the local variations in "transmission" may be meaningfully considered in terms of a body-interior vs. body-exterior differentiation.

If the body-image frame of reference is physiologically meaningful, it should permit one to formulate certain ideas and concepts concerning physiological reactivity which will have some support from past observations and formulations.

#### DEDUCTIONS FROM BODY-IMAGE SCHEMA

Three kinds of deductions concerning patterns of physiological reactivity may be made in terms of the body-image schema:

1. It would be assumed that past observations would have uncovered reaction patterns which are linked with body exterior in distinction to body interior. Further, one would anticipate that those who had intensively studied various sorts of body reactivity phenomena would have found the body-exterior vs. body-interior differentiation to be a meaningful theoretical framework.

2. When varied physiological measures are obtained from individuals who have psychosomatic symptoms involving the body interior, these measures should indicate high reactivity of body-interior sectors. One would also expect such individuals to show low reactivity of body-exterior areas. Conversely, those with body-exterior symptoms should manifest high reactivity of the body exterior. They would probably show, in addition, low reactivity of the body interior. This second prediction is derived from the idea that those who have exterior body symptoms are indicating the unusual importance of the barrier quality

of their body exterior, and thus their relatively greater tendency to maximize reactions in this sector of the body rather than in the body interior. But those with interior symptoms are assumed to assign less importance to the body exterior and more importance to the body interior.

3. Those who manifest unusually high physiological reactivity in the exterior body layers should have a more active and aspiring attitude toward life than those who do not show high reactivity in the body exterior. This prediction is based on the finding (14) that in a normal group those who emphasize the importance of their body exterior express a relatively high level of aspiration (as defined by TAT measures).

In connection with the first hypothesis, we have searched the literature widely for conceptual models developed by others to explain their observations of physiological reaction patterns which would correspond to the schema we have postulated. There are several such models worth noting.

Burton and Edholm (2) in their recent book *Man in a Cold Environment* report a diversity of data concerning the manner in which the body adjusts to sharp temperature changes, and how it maintains long-term temperature homeostasis. In the process of schematizing the various observed temperature-regulating mechanisms they arrived at a formulation analogous to the exterior-interior concept:

The temperature of the deep tissues, the brain, the heart and abdominal viscera such as the liver is not exactly the same. . . . These are close enough to uniformity and constancy to justify the simplifying concept of a central deep "core" of the body of uniform, regulated temperature, surrounded by a "shell" of cooler peripheral tissues, whose temperature moreover is dependent on that of the environment as well as on physiological factors. Indeed, the homeothermy of the "core" is accomplished, in great measure, by the adjustment of the temperatures in the "shell."



That is, Burton and Edholm have perceived, in the very complicated physiological processes involved in the body's adjustment to the changing temperatures of the surrounding environment, a specialization in function that correlates with body exterior vs. body interior.

Kleitman (20) has described a model, having to do with levels of consciousness, which is pertinent to the exterior-interior distinction. He has developed an "evolutionary theory of consciousness" which is based on a differentiation between the effects on consciousness of stimuli originating in the viscera and stimuli coming from more exterior sources, especially the proprioceptive end organs. He points out that in the newborn infant wakefulness (consciousness) occurs briefly every four or five hours, and is related to afferent impulses from the viscera. Similarly, he says of decorticate dogs that they "almost invariably moved about for a few minutes after being fed, apparently because of afferent impulses from the colon resulting from the gastrocolic reflex. They would then defecate and immediately go to sleep" (20, p. 81). This kind of consciousness is described as "wakefulness of necessity." It is a function primarily of afferent impulses from the internal viscera. Kleitman indicates that with the myelination of the afferent system in the child, the reactivity of the cortex to sensory stimuli increases. It becomes more capable of reacting to stimuli from without, particularly to impulses originating in the proprioceptive end organs. The development of the muscular system becomes a salient source of proprioceptive excitation which, in combination with other sensory impulses, leads to a higher level of consciousness. Kleitman designates this higher level of consciousness as "wakefulness of choice." For our purposes what is most interesting about this

sharper form of consciousness is that it is strongly linked with the increasing potency of impulses originating in an important part of the outer body layers (viz., the muscular-proprioceptive system). In summary of Kleitman's views, one may say that he associates a first, lower level of consciousness with the individual's inability to react to little more than the impulses originating in the viscera of the body interior. He regards later and higher forms of consciousness as a function of increasing reactivity to sensory stimuli from without, particularly those involved in proprioception.

Lorr, Rubenstein, and Jenkins (26) factor-analyzed ratings made by psychotherapists of 184 patients whom they were treating. These ratings concerned a wide range of behavior, including body complaints. The ratings of the body complaints clustered into three different factors. One factor was labeled a "factor of endodermal dysfunction," and involves complaints relating mainly to the upper and lower portions of the endodermal tube (e.g., gastrointestinal symptoms). A second factor involves complaints of "mesodermal origin." It includes mainly symptoms which are "anergic, respiratory, and cardiovascular." The third factor is of a "cerebrotonic" order, and particularly includes skin complaints. This pattern of clusters that emerged is obviously one which conceptualizes body complaints and body symptoms as occurring in layer-like fashion, with inner layers roughly distinguished from outer layers. It too shows certain analogies to the body-interior vs. body-exterior model.

Some of the more recent advances in our knowledge of the secretions produced by the adrenal medulla also lend support to the interior-exterior concept of body response. Gellhorn (17) indicates that it is now well established that the adrenal medulla secretes not only adrenalin but also nor-adrenalin.

Funkenstein (15) reviews data indicating support for the idea that there are specifically differentiated adrenalin-secreting and nor-adrenalin-secreting cells in the adrenal medulla. He refers also to a specific innervation of these cells, and to a differentiated hypothalamic representation for the corresponding secretory nerves. Adrenalin and nor-adrenalin produce rather different effects physiologically. Nor-adrenalin tends to result in moderate increase in systolic blood pressure, more marked increase in diastolic blood pressure, and a *decrease* in cardiac activity. However, adrenalin tends to produce only mild to moderate increase in diastolic pressure, but moderate to marked *increase* in cardiac activity. Adrenalin acts more forcefully in increasing cardiac output, whereas nor-adrenalin produces some of its most marked effects in terms of increased peripheral resistance. Apropos of this point, Goldenbert (as cited in [17]) found that intravenous nor-adrenalin led to increased blood pressure mainly as the result of increased peripheral resistance, whereas intravenous adrenalin stimulated a blood-pressure rise chiefly by means of an increased cardiac output. The accumulating data concerning adrenalin and nor-adrenalin suggest that their differential physiological effects fall into a pattern congruent with a body-exterior vs. body-interior differentiation.

More recently, Funkenstein, King, and Drolette (15) have demonstrated that a predominance of either the adrenalin or nor-adrenalin response in a stressful situation is correlated with certain patterns of behavioral response. They studied 69 students exposed to a stressful frustrating situation, and classified their reactions into three primary categories. The first category, "anger-out," was directed toward the experimenter who was producing the frustration. The second category, "anger-in,"

involved self-attack and self-criticism. The third category, "anxiety," involved feelings of apprehension and fright. Significantly, it was demonstrated that the anger-out group tends to show nor-adrenalin patterns of reaction. The anger-in and anxiety groups showed adrenalin patterns of response. Such data indicate that not only do adrenalin and nor-adrenalin produce physiological effects with a differential area of body response, but also that this differentiation is meaningfully linked with a difference in overt emotional behavior that is psychologically meaningful.

As noted later in this paper, the linkage of anger-out with the nor-adrenalin response, and the linkage of anger-in and anxiety with an adrenalin response, are just what one would anticipate from the body-image schema that has been proposed. That is to say, we have shown in our previous work that individuals who emphasize the body exterior, as contrasted to those who emphasize the body interior, have a higher level of aspiration and a more active, forceful approach to situations. It is therefore noteworthy that the subjects in the study carried out by Funkenstein and King who show a nor-adrenalin pattern of response (i.e., emphasis on peripheral reaction) are those who deal with the frustration directly by expressing their anger outwardly. However, the subjects manifesting the adrenalin pattern (where the response is relatively more interior) show less ability to deal with their feelings of frustration actively, and tend to express them more in self-criticism and generalized anxiety.

Kubie (21) has been very explicit in differentiating between the organs of the body, which he calls "organs which implement our relationships to the external environment" and "organs of internal economy." Those in the first group are "innervated predominantly by the somatomuscular and somatosensory ap-



paratus." They represent that part of the individual which is most consciously directed. However, the second grouping of organs "lies deep within the interior of the body and consequently are hidden from the individual's capacity for direct knowledge of himself. Furthermore, even though some of these organs have indirect connections with the outside world, our subjective awareness of them is absent or limited." Kubie associates disorders usually referred to as "conversion hysterias" with the "external" organ system, and he associates "organ neuroses" with the "internal" organ system.

Wenger (36), in the course of a factor analysis of a variety of physiological measures obtained from adults, found a factor which he designates as "representing certain functions of the skin and peripheral blood vessels." Similarly, Sanford (32) studied a group of children and adolescents with a battery of physiological and personality measures. His analysis of the intercorrelations among the physiological measures revealed "a syndrome—consisting of flushing, sweating, skin stroking intensity, odor, acne, and palpable thyroid." That is, there was a clustering of variables having mainly to do with skin appearance and skin responsivity.

The second deduction that was made from the body-image schema involved the relationship of site of psychosomatic symptom to pattern of physiological reactivity. This deduction finds some confirmation in past work. It should be specified at this point that, in order to keep the distinction between body exterior and body interior as clear as possible, measures of exterior body functioning will be restricted to GSR, muscle action potential, and direct methods for evaluating vasoconstriction and vasodilation of the peripheral blood vessels. Measures of interior body functioning will be restricted mainly to heart per-

formance and measures of changes in the stomach and colon.

Since intestinal disturbances and stomach ulcers are interior symptoms, one would predict, in terms of the exterior-interior body-image schema, that subjects with such symptoms would show heightened reactivity relative to body-interior measures. Karush, Hiatt, and Daniels (19) investigated a variety of physiological responses manifested by six ulcerative colitis patients during psychotherapeutic interviews. They obtained measures of colon reaction, salivary flow, and peripheral vascular changes. There was clear indication that specific stressful material elicited in the course of the therapy produced correlated colonic responses in the colitis patients. But the salivary responses and the vascular bed responses (exterior) did not show meaningful correlations with the stressful therapy material. These patients with an interior symptom showed their most specific physiological responses to stress in terms of an interior reaction.

Lewinsohn (24) obtained a series of physiological measures from patients with duodenal ulcers, patients with essential hypertension, psychiatric patients in whom muscular tension was the outstanding clinical symptom, and a control group of patients with hernia, hemorrhoids, and similar difficulties. One of his clearest findings, which he could not explain, was that the absolute skin resistance of the ulcer group was significantly greater than that of the control group. This relative inactivity of the skin layer (sweat glands) in the ulcer patients corresponds to the minimizing of the body exterior which one would predict in terms of the body-image schema. It is perhaps confusing that high skin resistance is equated with inactivity of the skin layer. However, this formulation is based on the assumption that lowered resistance is a func-



tion of the degree of secreting activity of the sweat glands. Lowered skin resistance represents active response of anatomical "units" in the skin layer. Van der Valk and Groen (35) have also found that skin resistance of patients with peptic ulcer is high.

Another of Lewinsohn's (24) most clear-cut findings was that patients with muscle-tension symptoms show more finger tremor under stress conditions than the control group, and are also characterized by a significantly lower heart rate than any of the other groups. These individuals whose clinical symptom is most obviously focused in the outer layers of the body (musculature) show a high level of reactivity in the musculature (exterior) and a low level of activity in terms of an interior measure. Little (25), in a study of peptic ulcer patients, found that they were higher in dermatographia persistence than a control group. That is to say, their peripheral vessels manifested less tone (and so may be assumed to have been less active) than did those of the control subjects. It should be acknowledged that, on the contrary, Wenger (40) found that a group of asthmatics (interior symptom) has significantly lower dermatographia persistence (i.e., higher vessel tone) than normals. Little (25) also demonstrated that ulcer patients had a significantly shorter heart period than a control group. In other words, as would be predicted, they showed a relatively higher rate of response with an interior organ.

There is another interesting finding reported by Little which points up the potential explanatory value of the exterior-interior concept. Little evaluated the effects of vagotomy on the physiological reactivity of a group of ulcer patients, and compared the shifts occurring from the pre- to the postoperative period with those changes in physiological reactivity occurring in a control

group of other surgical patients from the pre- to postoperative period. He compares the two groups in this fashion:

Both the control and experimental groups showed a drop in arterial pressure after operation, yet the mechanisms that produced this drop seem to be different in the two groups. In the controls, peripheral vasodilation as indicated by increased Finger Temperature and decreased Dermatographia Latency was the most relevant factor since the heart rate remained approximately constant. In the experimental group, however, there was an increased peripheral resistance, yet the blood pressure dropped in spite of it. In this group, the lowered heart rate is apparently the determining element. . . .

That is, following the cessation of the severe stress associated with the operation, both groups expressed their readjustment to the diminished stress by a drop in blood pressure. However, in the ulcer group the readjustment took place in terms of an interior organ, whereas in the control group the readjustment was more peripheral in character. This trend in the ulcer group is nicely congruent with the interior emphasis that has been postulated from the body-image schema.

The third basic deduction derived from the body-image schema focused on the idea that those who show unusually high physiological reactivity in the outer body layers will tend to be relatively high in level of aspiration, forcefulness, and drive to attain goals. Seymour (34) reports findings which are definitely in this direction. He indicates that children who show high GSR responsiveness (exterior) are significantly more interested in achievement and are more forceful in their behavior than children who are low in GSR reactivity. He further describes a study which demonstrated that adolescent boys who were perceived by their classmates as listless gave relatively low GSR responses to certain kinds of stimuli. In another phase of his data he found that

adolescent girls who were more daring and active in games gave larger GSR responses than those who were less so.

Cattell (4) describes a factor which is characterized by high responsivity in terms of one of our designated exterior measures (viz., GSR), and points out that this factor is psychologically associated with alertness, activity, and excitement. Low responsivity in this factor is correlated with sleepiness and passivity. Brown, as cited by Cattell (4), reports appreciable correlations of PGR with desire to excel. Cattell (4) describes a factor, "Nervous Disposition," which involves low PGR resistance and large PGR deflection; at the behavioral level it is linked with "alertness, hyperactivity." In the study by Sanford previously referred to (32), in which a cluster of measures prominently involving skin responsivity was discovered, it was also found that this cluster correlated positively with conscientious work and negatively with passive timidity. In a somewhat different approach to the physiological responsivity problem, Wishner (41, 42) has recently found that a physiological response cluster based on two exterior-type measures, GSR and muscle action potential, is associated with alertness. Such data tend to affirm the idea that high physiological reactivity of the body exterior is correlated with alertness, striving, and active expressiveness.

In general, it would appear that the over-all productiveness of the body-exterior vs. body-interior frame of reference has been sufficiently shown so that further exploration of the idea is justified. The next logical step, of course, is to set up a research design in which exterior and interior physiological measures are obtained from subjects who have been categorized in advance, relative to body-image indices. Such a study is being formulated and getting under way. If the schema is

supported by such research, there will then be available a frame of reference which can simultaneously embrace aspects of physiological reactivity phenomena, psychosomatic symptomatology, and personality defense reactions.

#### SUMMARY

We have discussed some of the problems involved in the categorization of physiological measures, and have emphasized particularly the lack of a physiological frame of reference from which to interpret patterns of physiological reactivity. In our previous work we have shown that the degree to which the individual conceives of his body boundaries as thick and armored vs. thin and permeable is significantly linked with site of psychosomatic symptomatology, and also with certain personality variables. We have suggested, on the basis of these findings, that patterns of physiological reactivity may be meaningfully viewed in terms of a contrast between body-exterior response and body-interior response. More specifically, it was postulated that those who, in their body image, emphasize the armored thickness of their body exterior would show relatively greater physiological reactivity in their exterior body layers (skin and muscle) than in their body interior. The converse of this was postulated to hold true for individuals who conceive of their body-image boundaries as permeable and easily penetrated. A range of literature has been reviewed which seems to lend support to such postulations.

#### REFERENCES

1. AX, A. F. The physiological differentiation of fear and anger in humans. *Psychosom. Med.*, 1953, 15, 432-442.
2. BURTON, A. C., & EDHOLM, O. G. *Man in a cold environment*. London: Edward Arnold, 1955.
3. CANNON, W. B. *Bodily changes in pain, hunger, fear and rage*. New York: Appleton-Century, 1915.



4. CATTELL, R. B. *Description and measurement of personality*. New York: World Book, 1946.
5. CLEGGHORN, R. A. The hypothalamus-endocrine system. *Psychosom. Med.*, 1955, 17, 367-376.
6. CLEVELAND, S. E., & FISHER, S. Behavior and unconscious fantasies of patients with rheumatoid arthritis. *Psychosom. Med.*, 1954, 16, 327-333.
7. CLEVELAND, S. E., & FISHER, S. Psychological factors in the neurodermatoses. *Psychosom. Med.*, in press.
8. DARLING, R. P. Autonomic action in relation to personality traits of children. *J. abnorm. soc. Psychol.*, 1949, 35, 246-260.
9. DARROW, C. W. Differences in the physiological reactions to sensory and ideational stimuli. *Psychol. Bull.*, 1929, 26, 180-201.
10. DARROW, C. W., & SOLOMON, A. P. Galvanic skin reflex and blood pressure in psychotic states. *Arch. Neurol. Psychiat.*, 1934, 32, 273-299.
11. DAVIS, R. C., BUCHWALD, A. M., & FRANKMANN, R. W. Autonomic and muscular responses, and their relation to simple stimuli. *Psychol. Monogr.*, 1955, 69, No. 20 (Whole No. 405).
12. EPPINGER, H., & HESS, L. *Vagotonia. Ment. nerv. Dis. Monogr.*, New York, 1915, No. 20.
13. FISHER, S., & CLEVELAND, S. E. The role of body image in psychosomatic symptom choice. *Psychol. Monogr.*, 1955, 69, No. 17 (Whole No. 402).
14. FISHER, S., & CLEVELAND, S. E. Body image and style of life. *J. abnorm. soc. Psychol.*, 1956, 52, 373-379.
15. FUNKENSTEIN, D. H., KING, S. H., & DROLLETTE, M. The direction of anger during a laboratory stress-inducing situation. *Psychosom. Med.*, 1954, 16, 404-413.
16. GANTT, W. H. *Experimental basis for neurotic behavior*. New York: Hoeber, 1944.
17. GELLHORN, E. *Physiological foundations of neurology and psychiatry*. Minneapolis: Univer. of Minnesota Press, 1953.
18. GREISMAN, S. E. The reaction of the capillary bed of the nailfold to the continuous intravenous infusion of levonor-epinephrine in patients with normal blood pressure and with essential hypertension. *J. clin. Invest.*, 1954, 33, 975-983.
19. KARUSH, A., Hlatt, R. B., & DANIELS, G. E. Psychophysiological correlations in ulcerative colitis. *Psychosom. Med.*, 1955, 17, 36-56.
20. KLEITMAN, N. *Sleep and wakefulness as alternating phases in the cycle of existence*. Chicago: Univer. of Chicago Press, 1939.
21. KUBIE, L. S. The central representation of the symbolic process in psychosomatic disorders. *Psychosom. Med.*, 1953, 15, 1-7.
22. LACEY, J. L. Individual differences in somatic response patterns. *J. comp. physiol. Psychol.*, 1950, 43, 338-350.
23. LACEY, J. L., & VAN LEHN, R. Differential emphasis in somatic response to stress. *Psychosom. Med.*, 1952, 14, 71-81.
24. LEWINSOHN, P. M. Some individual differences in physiological reactivity to stress. Unpublished doctor's dissertation, The Johns Hopkins Univer., 1954.
25. LITTLE, K. B. An investigation of autonomic balance in peptic ulcer patients. Unpublished doctor's dissertation, Univer. of California, 1950.
26. LORR, M., RUBENSTEIN, E., & JENKINS, R. L. A factor analysis of personality ratings of out-patients in psychotherapy. *J. abnorm. soc. Psychol.*, 1953, 48, 507-514.
27. MALMO, R. B., & SHAGASS, C. Physiologic studies of reaction to stress in anxiety and early schizophrenia. *Psychosom. Med.*, 1949, 11, 9-24.
28. MALMO, R. B., & SHAGASS, C. Physiologic study of symptom mechanisms in psychiatric patients under stress. *Psychosom. Med.*, 1949, 11, 29-39.
29. MCCLEARY, R. A. The nature of the galvanic skin response. *Psychol. Bull.*, 1950, 47, 97-117.
30. MOWRER, O. H. *Learning theory and personality dynamics*. New York: Ronald, 1950.
31. REYMERT, M. L. *Feelings and emotions*. New York: McGraw-Hill, 1950.
32. SANFORD, R. N., ADKINS, M. M., MULLER, R. B., & COBB, E. Physique, personality and scholarship. *Monogr. Soc. Res. Child Developm.*, 1943, 7, No. 34.
33. SCHAFER, H. R. Behavior under stress: a neurophysiological hypothesis. *Psychol. Rev.*, 1954, 61, 323-333.
34. SEYMOUR, R. B. Personality correlates of electrodermal resistance to response. Unpublished doctor's dissertation, Univer. of California (Berkeley), 1950.



35. VAN DER VALE, J. M., & GROEN, J. Electrical resistance of the skin during induced emotional stress. *Psychosom. Med.*, 1950, 12, 303-314.
36. WENGER, M. A. Measurement of individual differences in autonomic balance. *Psychosom. Med.*, 1941, 3, 427-434.
37. WENGER, M. S. Stability of measurement of autonomic balance. *Psychosom. Med.*, 1942, 4, 94-105.
38. WENGER, M. A., & ELLINGTON, M. Measurement of autonomic balance in children: method and normative data. *Psychosom. Med.*, 1943, 5, 241-253.
39. WENGER, M. A. Preliminary study of the significance of measures of autonomic balance. *Psychosom. Med.*, 1947, 11, 301-309.
40. WENGER, M. A. Studies of autonomic balance in Army Air Forces personnel. *Comp. psychol. Monogr.*, 1948, 19, No. 4 (Serial No. 101).
41. WISHNER, J. Neuroses and tension: an exploratory study of the relationship of physiological and Rorschach measures. *J. abnorm. soc. Psychol.*, 1953, 48, 253-260.
42. WISHNER, J. The concept of efficiency in psychological health and in psychopathology. *Psychol. Rev.*, 1955, 62, 69-80.
43. WOLFF, H. G. *Life stress and bodily disease*. Baltimore: Williams & Wilkins, 1950.
44. WOLFF, P., & LEVINE, J. Nocturnal gastric secretions of ulcer and nonulcer patients under stress. *Psychosom. Med.*, 1955, 17, 218-226.

(Received February 27, 1956)

## THE PHYSIOLOGICAL BASIS OF FORM PERCEPTION IN THE PERIPHERAL RETINA

R. H. DAY

*University of Sydney*

If the point of foveal fixation is held constant, and the angle subtended at the retina by that point and a stimulus object is decreased, either gradually with continuous exposure or by discrete steps with brief exposures, the appearance of the latter is perceived with progressively increasing clarity (4, 11, 12, 15, 17, 22, 30, 34, 35). There is now a large body of data concerned with qualitative and quantitative aspects of peripheral visual perception, but no attempt has yet been made to deal comprehensively with these data from a theoretical standpoint. Previous studies have been mainly of the fact-finding or exploratory variety, and have been little concerned with relating the subjective data to visual theory.

In an earlier paper (6) it was shown that the phenomena of foveal form perception, under the condition of increasing contrast between the stimulus figure and its surround, could be accounted for satisfactorily by the statistical theory of neurophysiological activity in the primary projection system of the eye. This paper suggested also that changes in nonfoveal form perception arising from changes in the retinal locus of the image might be dealt with by this theory. The aim of this contribution is to carry out a detailed analysis of these data within the framework of the statistical theory of visual perception.

No attempt will be made here to describe the nature of the statistical theory since this has already been outlined fully (6) and, together with the original statement of the theory (20), there are also comprehensive reviews of it in the psychological literature (23, 24, 31). It is sufficient to state here that this theory

employs the notion of populations of neural elements firing in a statistical fashion leading to a near-Gaussian distribution of excitation in area 17. Among the factors contributing to this distribution are reciprocally overlapping systems of fibers in the visual pathway giving rise to neural summation, neural recovery cycle characteristics, as well as diffractive properties, and nystagmatic movement of the eye. Visual acuity in its various forms is held to be a correlate of the peaking or sharpening of the cortical distribution of excitation. The sharper or more peaked is this distribution the greater is the resolution of fine detail. When the slope and height of the excitation gradient exceed certain minimum (threshold) values, the localization of an edge or contour in subjective visual space coincides with the location of maximal excitation in area 17. The clearness of a contour will depend upon the sharpness of the excitation peak and the steepness of the excitation gradient. In the case of an extended and flattened distribution, a line, for example, would appear broadened and blurred along its edges as well as tending to merge with the background.

### COMMENTS ON EXPERIMENTAL METHOD

Before proceeding to an analysis of the experimental data, it is as well to examine carefully the experimental techniques employed in this type of investigation, since these can affect considerably the validity of the results obtained.

Studies of form perception in the peripheral retina have used two methods. In the *moving* method the stimulus object is moved slowly but continu-

ously toward the fixation point. An objection raised against this method (34) concerns the tendency of the image to disappear with prolonged fixation (8, 9, 33). Disappearance of the image could lead to the loss of useful introspective data relating to the appearance of the stimulus figure at various removes from the fovea. In the *stationary* method the test object is momentarily exposed for a controlled time interval at different displacements from the point of fixation. Apart from the criticism of the moving method mentioned above, there appears to be no reason why these methods should not be regarded as giving essentially similar results. It is conceivable, however, that the mode of stimulation of the retinal "on-off type" receptors is not the same under the two conditions of exposure.

An important aspect of procedure which needs to be considered concerns the knowledge of the stimulus objects which *O* possesses before undergoing a trial, or the opportunity provided during a series of trials for *O* to become familiar with the stimulus objects. Most studies (4, 11, 12, 17, 22, 30, 34, 35) have either familiarized *Os* with the stimulus figures before beginning the experiment, or arranged the trials in such a manner that *Os* would grow very familiar with these test objects during the course of a series of trials. Collier actually took steps to ensure that *Os* "were trained in every type of reaction for which data would be later recorded" (4, p. 283). Using this procedure, it would be expected that by the completion of a series of trials *Os* would tend merely to recognize rather than "see" the stimulus figures. Drury has drawn attention to this difficulty:

It is difficult, if not impossible, to estimate the effects of recognizing, instead of "seeing," the known forms. To touch off a verbal name or other epithet, it is only necessary that the stimulus-pattern should be sufficient to suggest

"that form." Practice and repetition make *O* more and more expert in distinguishing the several members of a series (seen at times with full foveal adequacy). The experiment comes to be, therefore, more a process of coaching in the interpretation of indistinct signs than a discovery of retinal adequacy at various removes from the fovea (7, p. 639).

This criticism of a widely used experimental procedure is perhaps more pertinent in those experiments where the stimulus figure is eccentric to the foveal fixation point, than in those where the figure is viewed foveally under the condition of increasing contrast with respect to the surround. In the latter type of experiment *O* uses that part of the retina (fovea) which is not only efficient in the resolution of fine detail, but is well adapted for and regularly employed in form perception. Even though a stimulus object might be deficient in those qualities essential to the perception of a particular form, *O* can describe confidently what is perceived at a certain stage in the perceptual process. With peripheral presentation of the stimulus object, however, a region of the retina which is capable of perceiving, at the best, gross form characteristics only, is stimulated. Further, this part of the retina is seldom used in the perception and recognition of form qualities. It is to be expected, then, that in the typical campimetric experiment *O* would tend to rely even more upon what has been learned regarding characteristics of the stimulus figure. In short, because of the inefficiency in form perception of the peripheral perceiving mechanism, *Os* would be very ready to use as much as possible any knowledge acquired with respect to the nature of the stimulus figures. Where introspective data are reported fully (15), the evidence suggests that even without any knowledge of the stimulus figures *Os* tend to fit the figure to some preconceived form. It is unfortunate that no control has been exercised over this factor of preknowledge



of, or increasing familiarity with, the stimulus figure. The failure to do so has led to the exclusion from this analysis of otherwise valuable data.

#### PHYSIOLOGICAL PROCESSES

Reciprocal overlap of ascending neuron chains, the presence of lateral connecting neurons, and the convergence of receptor cells upon a single ganglion cell are some of the structural properties of the visual projection system which are of importance to the statistical theory in relation to peripheral vision.

Reciprocal (20) or partially shifted (18) overlap refers to a typical arrangement of cross connections between ascending fibers resulting in the divergence of impulses at one synaptic level and re-concentration of these at another synaptic relay higher in the system. The outcome of this convergence of impulses is spatial summation. Figure 1 shows how cross connections between a "core" fiber and its collaterals give rise to a spreading of excitation at one synapse ( $S_1$ ) but concentrate it at the next ( $S_2$ ). Activity in the collateral fibers at the level of  $S_2$  would tend to facilitate the core fiber, thus producing a differential rate of firing with consequent peaking of the cortical distribution of excitation. Spreading and concentration of excitation in this manner could be expected to occur both at the retinal level and higher in the system at the levels of the geniculate and cortex. Reciprocally overlapping fiber systems are more characteristic of the peripheral retina than of the foveal region, where a point projection of neurons is more closely approached. The manner in which cortical peaking occurs with foveal stimulation is more dependent upon neural recovery cycle behavior than on the neural structure described above (20).

It has been pointed out (20) that peaking of the cortical excitation pattern will occur only so long as all the later-

ally interconnected fibers potentially available for use are not operating. When all the available fibers in a group are activated fully, with no further recruitment possible, saturation occurs, i.e., summation reaches a ceiling and the peaked distribution is replaced by a plateau-type distribution. This process has been called occlusion (5) and, as Granit and Harper (12) have pointed out, the over-all effect will be determined by the balance of summation and occlusion. The addition of fresh fibers impinging upon an already maximally discharging neuron is ineffective, since the neuron is saturated. The nature of the distribution under conditions of partial and complete saturation is shown schematically in Fig. 1.

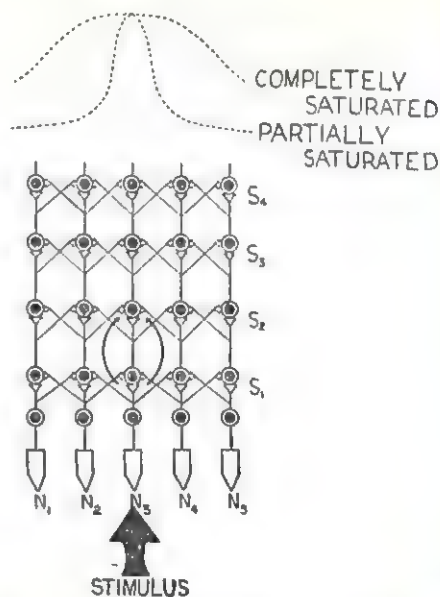


FIG. 1. Schematic representation of reciprocally overlapping fiber system.  $N_1, N_2, N_3, N_4$ , and  $N_5$  are parallel ascending fibers, and  $S_1, S_2, S_3$ , and  $S_4$  are four levels of synaptic junction. When a stimulus activates the "corefiber"  $N_3$  the resulting impulses tend to diverge at level  $S_1$  and to concentrate again at level  $S_2$ . The course of divergence and concentration of impulses is shown by the two arrows. At the top of the diagram are shown the probable shapes of the cortical excitation distributions under conditions of partial and complete saturation.

It has long been argued that the principal determinant of poor resolving power in the peripheral retina is the relatively greater ratio of receptor elements to optic nerve ganglia in this region (3, 28). When it is recalled that in parts of the peripheral retina the angular area from which receptor elements converge on to higher order neurons may be as great as  $2\frac{1}{2}$  degrees, it is plain that cellular convergence would interfere considerably with visual acuity.

Marshall and Talbot (20) have suggested that convergence is not incompatible with the peaking function of reciprocal overlap, since such acuity as is lost by convergence can be regained by overlapping systems of fibers at the level of the geniculate and cortical synapses. There is no reason to question this contention as a broad generalization, and, indeed, this is the point of view which will be adopted here. It is reasonable to suppose, however, that, because of a gradient of increasing convergence from fovea to periphery, the degree to which reciprocal overlap overcomes the effects of convergence will become less as the outer periphery is approached.

Chiewitz<sup>1</sup> (3) has shown that the ratio of receptors to inner nuclei increases regularly from fovea to periphery. This ratio is 1.09, 10.75, 42.00, and 80.00 to 1 for distances of 1.0, 3.2, 4.6, and 6.0 mm., respectively, from the fovea. Thus, whereas in regions near the fovea the low degree of cellular convergence would not greatly interfere with peaking arising from reciprocal overlap, the much higher ratio of receptor to inner cells in the outer periphery would severely limit this function. The actual manner in which the two neural processes consequent upon reciprocal

overlap and cellular convergence interact is by no means clear. It does seem, however, that whereas reciprocal overlap spreads and later concentrates streams of impulses from a small area, resulting finally in a peaked distribution, cellular convergence tends to bring together impulses from a wide area (1), resulting in summation at a single ganglion cell. Thus the effect of convergence will be to flatten and extend the cortical pattern of excitation, whereas reciprocal overlap will serve the function of sharpening the excitation peaks. The processes of convergence of impulses from a wide area upon a single cell, and of spreading and concentration of impulses due to reciprocal overlap, are not to be thought of as entirely separate neural functions operating independently of each other. These activities are doubtless intermingled, and retinal elements such as the internuncial probably serve both functions. What is of interest to the present analysis is the degree to which each process occurs relative to the other, for upon this will depend the shape and gradient of the cortical distribution of excitation.

The reciprocally overlapping system of fibers serving the function of cortical peaking, and the increasing convergence of receptor of ganglion cells from the fovea outwards limiting this process, would appear to be the basis of dyscric (26) or protopathic (16) and epicritic (26) vision. Certainly such a proposal fits more closely the facts than does an explanation in terms of the type and distribution of receptor cells. Parsons (26) has drawn attention to certain epicritic features of dyscric vision. That is, the difference between these two kinds of vision is a relative one only. The increasing limitation placed upon the peaking process from fovea to periphery by the gradually increasing receptor to ganglion convergence ratio would seem to be an acceptable physio-

<sup>1</sup> Chiewitz's investigations were carried out without the benefit of the chrom-silver staining methods. More recent investigations indicate, however, that the findings with respect to cellular convergence ratios are accurate.



logical explanation of the two kinds of visual function.

#### THEORETICAL INTERPRETATION OF EXPERIMENTAL DATA

*Qualitative data.* The data to be treated here are those reported by Zigler *et al.* (35) and Grindley (15): Inclusion of data from Zigler *et al.* requires some justification, since this work was criticized above on the grounds that the experimental conditions permitted Os to grow familiar with the series of stimulus objects during the course of the experiment. Use is made of this report here since the peripheral perceiving process has been divided into a number of stages, which facilitates this analysis. Also, results from Grindley's (15) investigation, to which the above criticism does not apply, are in general agreement with those of Zigler *et al.* A further study by Galli (11) reports data which correspond to those of Zigler *et al.*

Zigler *et al.* (35), using a wide variety of stimulus patterns and employing the moving method, have shown that the perceptual process can conveniently be divided into four stages as the image moves from the outer periphery toward the fovea. The first stage occurs when *O* first becomes aware of the total stimulus object, i.e., the stimulus figure and its background. This is referred to as the "figureless field" stage. Although *O* is aware of the presence of a stimulus object in the peripheral visual field, there is no differentiation between figure and ground. The second stage, that of the "formless figure," is reached when *O* reports that he is aware of "something" upon the background, i.e., when differentiation into figure and ground first occurs. Initially, *O* usually reports the presence of a faint spot or patch which in open figures is reported as being "filled in" or solid. The third stage is that of the "form-like figure" beginning at the point at which any suggestion in

regard to shape is reported. Initially the detail of the figure becomes organized into a figure which is essentially unclassified. Only the most general characteristics of shape can be described. These include such features as "straight-sided," "rounded," and "elongated" processes. The final stage is reached when the figure is accurately perceived and correctly described or named.

As the image of the stimulus object traverses the retina from periphery to fovea, it moves from a region where there is a higher ratio of receptor to inner cells to one where this ratio is very much lower. Thus the peaking of cortical excitation consequent upon reciprocal overlap will tend to grow more marked as the fovea is more closely approached. With progressive change in the cortical pattern of excitation, from an extended and flattened distribution with a slight gradient to one which is sharpened and the gradient steep, so, subjectively, edges will become better defined and detail clearer. In the parafoveal and foveal regions, peaking becomes more pronounced as a result of the neural recovery cycle behavior of the fiber groups involved. (For a full explanation of this process see [20].) In terms of the nature of the cortical excitation pattern, the four stages of perception defined by Zigler *et al.* (35) are neurophysiologically represented by a progressive sharpening of cortical excitation peaks and increases in the steepness in the slopes of the excitation gradients. This process is represented in Fig. 2, in which is shown the probable nature of the cortical pattern of firing at the four perceptual stages described above.

Grindley (15), using the stationary method, reports a number of "errors" of perception. Among these are the perception of an inverted T figure as a triangle with apex uppermost, two squares with corners superimposed seen



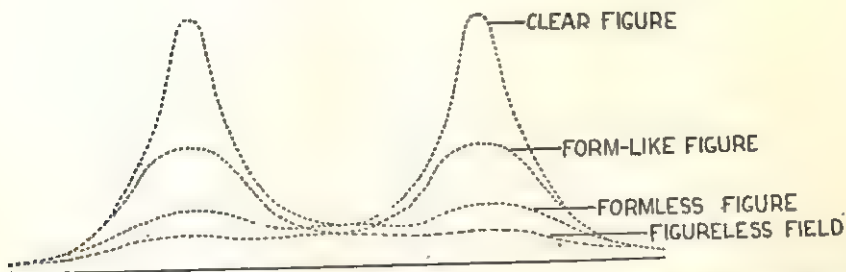


FIG. 2. Representation of cortical excitation distributions about the opposite edges of a figure for four stages of peripheral form perception. The slight increase in the height of the distributions due to facilitation within the area of overlap is shown for the "figureless field" and "formless figure" stages.

as a cross within a circle, and an open circle seen as a solid figure. Each of these phenomena can be attributed to a filling-in effect (6), correlated neurophysiologically with an extended and flattened cortical-excitation pattern. In the absence of cortical peaking an edge will appear as blurred and extended, as well as tending to merge with the background. In the case of the open circle seen as a solid circle or a disc, the extended distribution of excitation about the inner edge of one part of the circle will be continuous with the extended distribution about that part of the edge opposite to it. Figure 2 shows how, because of the flattened distribution of slight excitation gradient, this would occur. Another phenomenon of peripheral perception reported by Grindley (15) is the appearance of an open figure as an open figure but with a dark patch in the central part of the figure. Since the two cortical "ridges" of excitation about the opposite sides of a figure would tend in some cases to overlap, it is likely that those groups of fibers involved within the area of overlap would tend to facilitate each other. The outcome of such facilitation would be another peak or "ridge" of excitation at the center of the figure, although no edge or contour exists. Depending upon the size of the figure and the retinal locus of the image, this effect would lead to above-threshold excitation, thus giv-

ing the subjective impression of a patch at the center of the stimulus figure. Such a patch could occur alone or surrounded by the blurred contour of the figure, depending upon the heights of the three distributions. This process is shown diagrammatically in Fig. 2.

Blurring of the outline of a figure and the tendency for edges and contours to merge with background in peripheral vision would tend to obscure the shape of a figure, so that it would appear as a more or less circular patch or spot. This would be the case at the formless figure stage. Thus the two superimposed squares would be perceived as roughly circular, with the superimposed corners of the two squares forming a blurred cross within the circle. The blurring and filling-in effect would be responsible for an inverted T figure being perceived as a triangle with apex uppermost.

It must be stated, however, that there are a number of perceptual phenomena in peripheral vision which cannot be interpreted entirely in terms of the processes which have been outlined. If further determinants in peripheral visual perception—such as various types of eye movement (physiological nystagmus, drift, saccades), loss of fixation, disappearance of peripheral image, and O's preconceived notions of figure qualities—are also considered, then most of the phenomena are explicable.

*Effect of stimulus-object size.* Most investigators (15, 17, 30) are agreed that large stimulus figures are clearly seen and accurately described at a greater degree of eccentricity than are smaller figures. At a given retinal locus a small open circle would appear filled in, since the excitation distributions representing opposite edges would overlap, thus giving the subjective appearance of solidity. With larger circles, however, although the outline would still appear blurred and hazy, the circle would appear open, since the two cortical distributions would not overlap, or would overlap to the same extent as in the case of the smaller circle. In the same manner, gross aspects of form would not be obscured to the same extent as finer detail. In this respect it can be hypothesized that the degree of eccentricity from the fixation point at which a figure can be perceived clearly will depend upon the perimeter-to-area (P/A) ratio (2). This ratio defines the amount of edge detail relative to the area of the figure.

*Effects of exposure time.* In an examination of the effects of exposure time on peripheral visual perception, Grindley (15) varied exposure time from 0.01 to 1.00 sec., with the stimulus located at angular distances of 20, 40, and 60 degrees from the fixation point. With very short exposures the pattern was either not seen at all, or, when it was seen, it appeared as a blurred patch. With longer exposures the pattern was reported as blurred throughout the time for which it was exposed, clear at first, then growing blurred; or blurred initially and clear later. A report by Meisenheimer (21) is in agreement with these results. Unfortunately Grindley's study does not include detailed results for the three eccentric angles used. Nevertheless, the general findings for variation in exposure time are what

would be expected on the basis of the statistical theory.

The nonappearance of the stimulus object, or its appearance as a blurred patch, can again be attributed to a flattened distribution of cortical activity about the cortical representation of edges and contours. In the case of nonappearance of the stimulus pattern, the height of this distribution would be below the threshold for visual experience, and with a blurred patch slightly above that threshold. For each of these brief exposure phenomena the absence of peaking can be attributed not only to the inhibition of the mechanism of reciprocal overlap by cellular convergence, but also to an insufficiency of time for the peaked distribution to develop. Evidence for this is provided by reaction-time studies, which show that in the periphery RT increases from 0.004 sec. at 3 degrees to 0.024 sec. at 45 degrees (19, 27) indicating an increase in neural transmission time and synaptic delay. Even in the dark adapted state, when the peripheral receptors are in a very much more sensitive state, increase in reaction time compared with the fovea is insignificant (27). In the inner periphery the insufficiency of time for a peaked distribution to build up would be the principal determinant of the nonappearance of the stimulus object. In the outer periphery this time factor would operate along with the inhibitory effects of convergence.

The brief appearance of the figure at the beginning of the exposure, followed by blurring for the remainder of the exposure, is probably attributable to a saturation process already described. This process is described more fully in the next section dealing with "fixation blindness." Depending upon the retinal locus of the image and the exposure time involved, initial blurring followed by a clear image is due, probably, to the time necessary for the building up of



cortical peaks of activity. It would be expected that this phenomenon would be more characteristic of the outer peripheral region.

Whether the stimulus pattern is blurred throughout the time of exposure, appears clear at first, and then fades to become blurred or to disappear completely, or vice versa, will be a complex function of other factors as well as those relating to retinal and cortical processes. These include retinal locus of the image, stimulus size, shape, and intensity, together with constancy of fixation, knowledge concerning the stimulus pattern, and eye movements. So far, the conditions under which an object is clear or blurred, or passes from one of these states to the other, have not been specified. There is obviously a considerable need for experimentation to determine under what conditions these various phenomena occur.

*Disappearance of the peripheral image.* Fading of the peripheral image, first reported by Troxler (33), and sometimes referred to as "fixation blindness" (14), occurs usually after two to three seconds of fixation, and is limited to the peripheral retina. This effect occurs more rapidly at large peripheral angles and with greater stimulus intensities, but less rapidly with large stimulus objects. Restoration of the image can be brought about by brief loss of fixation, or by rapid removal and replacement of the stimulus figure. Dunlap (9) reported that reappearance of the image occurs spontaneously, sometimes in the form of the stimulus object itself, sometimes as a darker spot, and occasionally as a coronal effect surrounding the position where the stimulus object was last seen. Dunlap reports also that: "A haze or blur of diffused light was discernible, covering a considerable part of the visual field around the spot, and this haze of light

remained after the brighter spot had disappeared" (9, p. 204).

Discrimination between a figure and its background is dependent upon the perception of an edge or contour separating the two. The cortical correlate of the clear perception of an edge is a sharpened peak to the cortical excitation pattern and a steep excitation gradient. In the peripheral retina, as has been pointed out, this sharpening is dependent upon a reciprocally overlapping system of fibers. When such a system becomes saturated the cortical excitation distribution becomes flattened, and the excitation gradient less steep (Fig. 1). The subjective outcome of saturation would be the blurring and obscuring of edges and borders, with consequent loss of clear differentiation between figure and ground. Indeed Dunlap's (9) description of "a haze or blur of diffused light" is precisely how an edge would be expected to appear under conditions of saturation.

A process of saturation in the manner suggested is well supported by experimental evidence. The rapid onset of the process and a negatively accelerated course (14) is to be expected from the structural arrangement of fibers responsible for peaking. A more intense stimulus would result in more rapid saturation of the group of fibers involved, and a large stimulus object would be less likely to be obscured by edge blurring and haziness than would a smaller one. Brief loss of fixation or removal of the stimulus would permit the saturated fiber group to recover, and so restore the image.

Spontaneous reappearance of the image is more difficult to account for. Figure 1 shows, however, that, although the cortical excitation pattern under conditions of saturation is flattened and the gradient slight, the height of the distribution remains above threshold for perception. That is to say, although con-



tour is obscured by blurring about the edges, resulting in virtual loss of the figure, the height of the distribution remains at threshold level. Assuming nystagmatic eye movement and drift, however, further fibers will be activated, causing the point of maximal cortical excitation to shift, and thus resulting in the brief reappearances of the image before these, too, become saturated. Since the distribution has remained above threshold, activation of further fibers, resulting in the definition of edges, would readily result in the reappearance of the image.

*An experiment.* The theory of peripheral form perception which has been outlined possesses the advantage of permitting relatively precise predictions to be made regarding the changes in the appearance of a stimulus pattern as the distance of the image from the fovea is decreased. An as yet unpublished investigation carried out by the writer was designed to test some of the predictions and is worth quoting in this context.

The six figures used are shown in Fig. 3. These were presented by means of an exposure device mounted on a Stoelting campimeter. The Os were quite unfamiliar with the figures, and no comment was made by *E* after a description of a particular test object had been offered. The Os both described and drew what they saw at each angle of presentation.

Generally it was predicted that in the outer periphery the figures would be seen as dark, almost formless, patches on a whitish ground. Closer to the fovea, small edge indentations would be filled in and corners rounded by blurring at the edges. White spaces enclosed within the figure would tend to be seen

as continuous with the rest of the figure, due to the filling-in effect. Forms 2, 5, and 6 would be seen at some stage as rectangles before the emergence of detail. Forms 1 and 3 would be seen as roughly triangular, and Form 4 seen first as a hazy rounded patch, then as diamond-shaped and finally as a blurred cross. In all cases these predictions have been correct. For example, the inverted L shape was seen first by one O at 70 degrees as a dark blur with a flattened top, at 50 degrees as a dark triangle with apex downward, and at 30 degrees as a boomerang shape. Again, Form 6 was perceived first as a blurred dark patch, and at between 50 and 30 degrees as a hollow square before it emerged clearly.

#### SUMMARY AND CONCLUSIONS

A theory of form perception in the peripheral retina, based upon the statistical theory of neurophysiological activity in the visual primary projection system, has been presented. The basic assumption in this theory is that an edge, border or contour has as its cortical correlate a near-Gaussian distribution of excitation in area 17. The sharpness of the peak of excitation of this distribution, and the slope of the gradient, determine the clarity with which form is perceived.

Certain methodological problems relating to the familiarity of O with the stimulus objects have been examined. Attention is drawn to the tendency of O merely to recognize and designate, as opposed to "seeing" and describing, the stimulus object when the experimental conditions are such as to allow O to become familiar with the stimulus objects.

It is suggested that in the peripheral retina the reciprocally overlapping systems of fibers give rise to a peaking of the cortical excitation pattern, whereas convergence of the receptor layer upon inner retinal cells tends to limit or in-

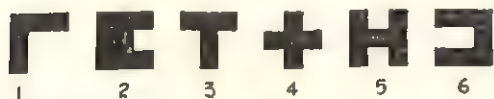


FIG. 3. Six forms used in experiment.

hibit this process. Since convergence increases from the fovea to the outer periphery, the latter is much less efficient in form perception. It is proposed that the processes consequent upon reciprocal overlap and cellular convergence form the basis of what has been called dyscritic and epicritic vision.

Data drawn from various sources have been examined and interpreted as evidence for the theoretical proposals. These data include qualitative aspects of peripheral form perception, effects of stimulus-object size and exposure time, and peripheral "fixation blindness."

An experiment designed to test some predictions of this theory is described briefly.

The statistical theory has now been evoked to deal with such diverse visual phenomena as visual acuity (20, 32), figural aftereffects (24), movement perception (23), and foveal form perception (6). With each of these phenomena the theory has dealt satisfactorily with the greater part of the experimental data. It is possibly pertinent to add that a not dissimilar theory has been proposed in regard to somesthetic sensitivity (29).

Because knowledge of the neurophysiological processes of the visual system is still incomplete, it is inevitable that many of the theoretical proposals are speculative. The theory does have, however, the important virtue of permitting readily testable hypotheses to be set up and subjected to physiological and psychological experiment. Until this is done, the validity of the statistical theory, and such modifications that may prove necessary to increase its validity, must remain uncertain.

#### REFERENCES

1. BEITEL, R. J. Spatial summation of subliminal stimuli in the retina of the human eye. *J. gen. Psychol.*, 1934, 10, 311-327.
2. BITTERMAN, M. E., KRAUSKOPF, J., & HOCHBERG, J. E. Threshold for visual form: a diffusion model. *Amer. J. Psychol.*, 1954, 67, 205-219.
3. CHIEWITZ, J. H. Untersuchungen über die Area centralis retinae. *Arch. f. Anat. u. Entwicklungsgesch.*, 1889, Suppl. 139.
4. COLLIER, R. M. An experimental study of form perception in peripheral vision. *J. comp. Psychol.*, 1931, 11, 281-289.
5. COOPER, S., DENNY-BROWN, D. E., & SHERRINGTON, C. S. Interaction between ipsilateral spinal reflexes acting on the flexor muscles of the hind limb. *Proc. Roy. Soc., Lond.*, 1927, 101, 262-303.
6. DAY, R. H. Application of the statistical hypothesis to form perception. *Psychol. Rev.*, 1956, 63, 139-148.
7. DRURY, M. B. Progressive changes in non-foveal perception of line patterns. *Amer. J. Psychol.*, 1933, 45, 628-646.
8. DUNLAP, K. Die Wirkung gleichzeitiger Reizung von zentralen und exzentrischen Netzhautstellen. *Arch. f. d. ges. Psychol.*, 1912, 24, 299-304.
9. DUNLAP, K. Light-spot adaptation. *Amer. J. Physiol.*, 1921, 55, 201-211.
10. FREEMAN, E. Anomalies of peripheral visual acuity. *J. exp. Psychol.*, 1929, 12, 324-340.
11. GALLI, A. La percezione della forma nella visione periferica. *Pubbl. Univer. Catol. d. S. Cuore*, 1931, 6, 1-27.
12. GEISSLER, L. R. Form perception in indirect vision. *Psychol. Bull.*, 1926, 23, 135-136.
13. GRANIT, R., & HARPER, P. Comparative studies in the peripheral and central retina. II. Synaptic reactions in the eye. *Amer. J. Physiol.*, 1930, 95, 211-228.
14. GRANIT, R., & VON AMMON, W. Comparative studies on the peripheral retina. III. Some aspects of local adaptation. *Amer. J. Physiol.*, 1930, 95, 229-241.
15. GRINDLEY, G. C. Psychological factors in peripheral vision. *Med. Res. Council, Spec. Rep. Ser.*, 1931, No. 163.
16. HEAD, H. *Studies in neurology*. London: Frowde, Hodder & Stoughton, 1920.
17. KLEITMAN, N., & BLIER, Z. A. Color and form discrimination in the periphery of the retina. *Amer. J. Physiol.*, 1928, 85, 178-190.
18. LORENTE DE NÓ, R. Continuation of the study of the ammonic system. *J. Psychol. Neurol., Lbs.*, 1934, 46, 113-177.
19. LEMMON, V. W., & GEISINGER, S. M. Reaction time to retinal stimulation under light and dark adaptation. *Amer. J. Psychol.*, 1936, 48, 140-142.

20. MARSHALL, W. H., & TALBOT, S. A. Recent evidence for neural mechanisms in vision leading to a general theory of sensory acuity. In H. Klüver (Ed.), *Visual mechanisms*. Lancaster, Pa.: Jaques Cattell, 1942. Pp. 117-164.
21. MEISENHEIMER, J. Experimente im peripheren Sehen von Gestalten. *Arch. f. d. ges. Psychol.*, 1929, 67, 1-130.
22. MUNN, N. L., & GIEL, G. A. A note on peripheral form discrimination under dark adaptation. *J. gen. Psychol.*, 1931, 5, 78-87.
23. OSGOOD, C. E. *Method and theory in experimental psychology*. New York: Oxford Univer. Press, 1953.
24. OSGOOD, C. E., & HEYER, A. W. A new interpretation of figural after-effects. *Psychol. Rev.*, 1952, 59, 98-118.
25. ØSTERBERG, G. Topography of the layer of rods and cones in the human retina. *Acta Opthal. Suppl.*, 1935, 61, 1-102.
26. PARSONS, J. H. *An introduction to the theory of perception*. Cambridge: Cambridge Univer. Press, 1927.
27. POFFENBERGER, A. T. Reaction time to retinal stimulation, with special reference to the time lost in conduction through nerve centers. *Arch. Psychol., N. Y.*, 1912, No. 23, 1-73.
28. POLYAK, S. L. *The retina*. Chicago: Chicago Univer. Press, 1941.
29. RUCH, T. C. Sensory mechanisms. In S. S. Stevens (Ed.), *Handbook of experimental psychology*. New York: Wiley, 1951.
30. SALAMAN, M. Some experiments on peripheral vision. *Med. Res. Council, Spec. Rep. Ser.*, 1929, 6.
31. SENDERS, VIRGINIA L. The physiological basis of visual acuity. *Psychol. Bull.*, 1948, 45, 465-490.
32. SENDERS, VIRGINIA L. Visual resolution with periodically interrupted light. *J. exp. Psychol.*, 1949, 39, 453-465.
33. TROXLER, D. Ueber das verschwinden gegebener Gegenstände innerhalb unseres Gesichtskreises. *Opthal. Biblio*, 1804, 2, 1-53. (Quoted by J. W. Baird, The color sensitivity of the retina. *Publ. Carnegie Inst. Wash.*, 1905, No. 29.)
34. WHITMER, C. A. Peripheral form discrimination under dark adaptation. *J. gen. Psychol.*, 1933, 9, 405-419.
35. ZIGLER, M. J., COOK, B., MILLER, D., & WEMPLE, L. The perception of form in peripheral vision. *Amer. J. Psychol.*, 1930, 42, 246-259.

(Received March 23, 1956)



## INTERFERENCE AND FORGETTING

BENTON J. UNDERWOOD

*Northwestern University*<sup>1, 2</sup>

I know of no one who seriously maintains that interference among tasks is of no consequence in the production of forgetting. Whether forgetting is conceptualized at a strict psychological level or at a neural level (e.g., neural memory trace), some provision is made for interference to account for at least some of the measured forgetting. The many studies on retroactive inhibition are probably responsible for this general agreement that interference among tasks must produce a sizable proportion of forgetting. By introducing an interpolated interfering task very marked decrements in recall can be produced in a few minutes in the laboratory. But there is a second generalization which has resulted from these studies, namely, that most forgetting must be a function of the learning of tasks which interfere with that which has already been learned (19). Thus, if a single task is learned in the laboratory and retention measured after a week, the loss has been attributed to the interference from activities learned outside the laboratory during the week. It is this generalization with which I am concerned in the initial portions of this paper.

Now, I cannot deny the data which show large amounts of forgetting produced by an interpolated list in a few minutes in the laboratory. Nor do I deny that this loss may be attributed to interference. But I will try to show

that use of retroactive inhibition as a paradigm of forgetting (via interference) may be seriously questioned. To be more specific: if a subject learns a single task, such as a list of words, and retention of this task is measured after a day, a week, or a month, I will try to show that very little of the forgetting can be attributed to an interfering task learned outside the laboratory during the retention interval. Before pursuing this further, I must make some general comments by way of preparation.

Whether we like it or not, the experimental study of forgetting has been largely dominated by the Ebbinghaus tradition, both in terms of methods and materials used. I do not think this is due to sheer perversity on the part of several generations of scientists interested in forgetting. It may be noted that much of our elementary knowledge can be obtained only by rote learning. To work with rote learning does not mean that we are thereby not concerning ourselves with phenomena that have no counterparts outside the laboratory. Furthermore, the investigation of these phenomena can be handled by methods which are acceptable to a science. As is well known, there are periodic verbal revolts against the Ebbinghaus tradition (e.g., 2, 15, 22). But for some reason nothing much ever happens in the laboratory as a consequence of these revolts. I mention these matters neither by way of apology nor of justification for having done some research in rote learning, but for two other reasons. First, it may very well be true, as some have suggested (e.g., 22), that studies of memory in the Ebbinghaus tradition are not getting at all of the important

<sup>1</sup> Address of the president, Midwestern Psychological Association, St. Louis, Missouri, May, 1956.

<sup>2</sup> Most of the data from my own research referred to in this paper were obtained from work done under Contract N7 onr-45008, Project NR 154-057, between Northwestern University and The Office of Naval Research.

phenomena of memory. I think the same statement—that research has not got at all of the important processes—could be made about all areas in psychology; so that the criticism (even if just) should not be indigenous to the study of memory. Science does not deal at will with all natural events. Science deals with natural events only when ingenuity in developing methods and techniques of measurement allow these events to be brought within the scope of science. If, therefore, the studies of memory which meet scientific acceptability do not tap all-important memorial processes, all I can say is that this is the state of the science in the area at the moment. Secondly, because the bulk of the systematic data on forgetting has been obtained on rote-learned tasks, I must of necessity use such data in discussing interference and forgetting.

Returning to the experimental situation, let me again put in concrete form the problem with which I first wish to deal. A subject learns a single task, such as a list of syllables, nouns, or adjectives. After an interval of time, say, 24 hours, his retention of this list is measured. The explanatory problem is what is responsible for the forgetting which commonly occurs over the 24 hours. As indicated earlier, the studies of retroactive inhibition led to the theoretical generalization that this forgetting was due largely to interference from other tasks learned during the 24-hour retention interval. McGeoch (20) came to this conclusion, his last such statement being made in 1942. I would, therefore, like to look at the data which were available to McGeoch and others interested in this matter. I must repeat that the kind of data with which I am concerned is the retention of a list without formal interpolated learning introduced. The interval of retention with which I am going to deal in this, and several subsequent analyses, is 24 hours.

First, of course, Ebbinghaus' data were available and in a sense served as the reference point for many subsequent investigations. In terms of percentage saved in relearning, Ebbinghaus showed about 65 per cent loss over 24 hours (7). In terms of recall after 24 hours, the following studies are representative of the amount forgotten: Youtz, 88 per cent loss (37); Luh, 82 per cent (18); Krueger, 74 per cent (16); Hovland, 78 per cent (11); Cheng, 65 per cent and 84 per cent (6); Lester, 65 per cent (17). Let us assume as a rough average of these studies that 75 per cent forgetting was measured over 24 hours. In all of these studies the list was learned to one perfect trial. The percentage values were derived by dividing the total number of items in the list into the number lost and changing to a percentage. Thus, on the average in these studies, if the subject learned a 12-item list and recalled three of these items after 24 hours, nine items (75 per cent) were forgotten.

The theory of interference as advanced by McGeoch, and so far as I know never seriously challenged, was that during the 24-hour interval subjects learned something outside the laboratory which interfered with the list learned in the laboratory. Most of the materials involved in the investigations cited above were nonsense syllables, and the subjects were college students. While realizing that I am viewing these results in the light of data which McGeoch and others did not have available, it seems to me to be an incredible stretch of an interference hypothesis to hold that this 75 per cent forgetting was caused by something which the subjects learned outside the laboratory during the 24-hour interval. Even if we agree with some educators that much of what we teach our students in college is nonsense, it does not seem to be the kind



of learning that would interfere with nonsense syllables.

If, however, this forgetting was not due to interference from tasks learned outside the laboratory during the retention interval, to what was it due? I shall try to show that most of this forgetting was indeed produced by interference—not from tasks learned outside the laboratory, but from tasks learned previously in the laboratory. Following this I will show that when interference from laboratory tasks is removed, the amount of forgetting which occurs is relatively quite small. It then becomes more plausible that this amount could be produced by interference from tasks learned outside the laboratory, although, as I shall also point out, the interference very likely comes from prior, not interpolated, learning.

In 1950 a study was published by Mrs. Greenberg and myself (10) on retention as a function of stage of practice. The orientation for this study was crassly empirical; we simply wanted to know if subjects learn how to recall in the same sense that they learn how to learn. In the conditions with which I am concerned, naive subjects learned a list of ten paired adjectives to a criterion of eight out of ten correct on a single trial. Forty-eight hours later this list was recalled. On the following day, these same subjects learned a new list to the same criterion and recalled it after 48 hours. This continued for two additional lists, so that the subjects had learned and recalled four lists, but the learning and recall of each list was complete before another list was learned. There was low similarity among these lists as far as conventional symptoms of similarity are concerned. No words were repeated and no obvious similarities existed, except for the fact that they were all adjectives and a certain amount of similarity among prefixes, suffixes, and so on must inevitably occur. The

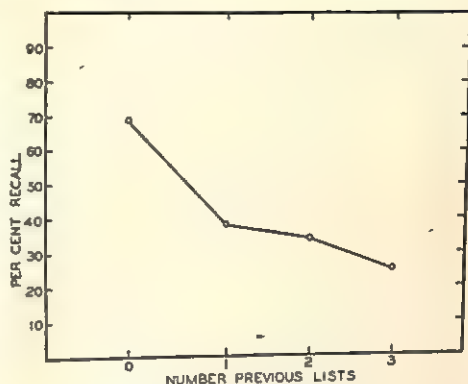


FIG. 1. Recall of paired adjectives as a function of number of previous lists learned (10).

recall of these four successive lists is shown in Fig. 1.

As can be seen, the more lists that are learned, the poorer the recall, from 69 per cent recall of the first list to 25 per cent recall of the fourth list. In examining errors at recall, we found a sufficient number of intrusion responses from previous lists to lead us to suggest that the increasing decrements in recall were a function of proactive interference from previous lists. And, while we pointed out that these results had implications for the design of experiments on retention, the relevance to an interference theory of forgetting was not mentioned.

Dr. E. J. Archer has made available to me certain data from an experiment which still is in progress and which deals with this issue. Subjects learned lists of 12 serial adjectives to one perfect trial and recalled them after 24 hours. The recall of a list always took place prior to learning the next list. The results for nine successive lists are shown in Fig. 2. Let me say again that there is no laboratory activity during the 24-hour interval; the subject learns a list, is dismissed from the laboratory, and returns after 24 hours to recall the list. The percentage of recall falls from 71 per cent for the first list to 27 per cent for the ninth.

In summarizing the more classical data on retention above, I indicated that



a rough estimate showed that after 24 hours 75 per cent forgetting took place, or recall was about 25 per cent correct. In viewing these values in the light of Greenberg's and Archer's findings, the conclusion seemed inescapable that the classical studies must have been dealing with subjects who had learned many lists. That is to say, the subjects must have served in many conditions by use of counterbalancing and repeated cycles. To check on this I have made a search of the literature on the studies of retention to see if systematic data could be compiled on this matter. Preliminary work led me to establish certain criteria for inclusion in the summary to be presented. First, because degree of learning is such an important variable, I have included only those studies in which degree of learning was one perfect recitation of the list. Second, I have included only studies in which retention was measured after 24 hours. Third, I have included only studies in which recall measures were given. (Relearning measures add complexities with which I do not wish to deal in this paper.) Fourth, the summary includes only material learned by relatively massed practice. Finally, if an investigator had two or more conditions which met these criteria, I averaged the values presentation

in this paper. Except for these restrictions, I have used all studies I found (with an exception to be noted later), although I do not pretend to have made an exhaustive search. From each of these studies I got two facts: first, the percentage recalled after 24 hours, and second, the average number of previous lists the subjects had learned before learning the list on which recall after 24 hours was taken. Thus, if a subject had served in five experimental conditions via counterbalancing, and had been given two practice lists, the average number of lists learned before learning the list for which I tabulated the recall was four. This does not take into account any previous experiments in rote learning in which the subject might have served.

For each of these studies the two facts, average number of previous lists learned and percentage of recall, are related as in Fig. 3. For example, consider the study by Youtz. This study was concerned with Jost's law, and had several degrees of learning, several lengths of retention interval, and the subjects served in two cycles. Actually, there were 15 experimental conditions and each subject was given each condition twice. Also, each subject learned six practice lists before starting the experimental conditions. Among the 15 conditions was one in which the learning of the syllables was carried to one perfect recitation and recall was taken after 24 hours. It is this particular condition in which I am interested. On the average, this condition would have been given at the time when the subject had learned six practice lists and 15 experimental lists, for a total of 21 previous lists.

The studies included in Fig. 3 have several different kinds of materials, from geometric forms to nonsense syllables to nouns; they include both paired-associate and serial presentation, with differ-

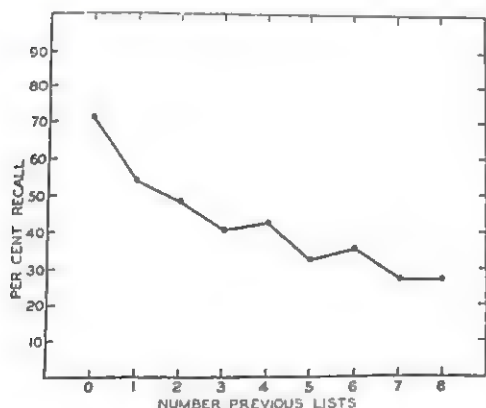


FIG. 2. Recall of serial adjective lists as a function of number of previous lists learned. Unpublished data, courtesy of Dr. E. J. Archer.

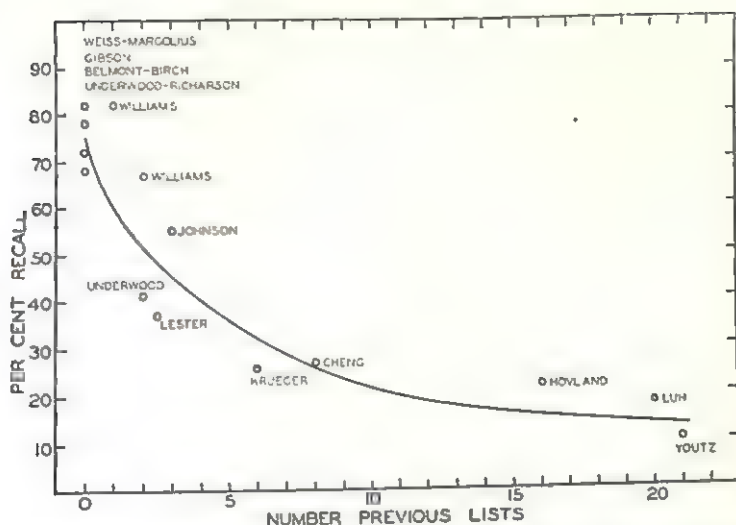


FIG. 3. Recall as a function of number of previous lists learned as determined from a number of studies. From left to right: Weiss and Margolius (35), Gibson (9), Belmont and Birch (3), Underwood and Richardson (33), Williams (36), Underwood (27, 28, 29, 30), Lester (17), Johnson (14), Krueger (16), Cheng (6), Hovland (11), Luh (18), Youtz (37).

ent speeds of presentation and different lengths of lists. But I think the general relationship is clear. The greater the number of previous lists learned the greater the forgetting. I interpret this to mean that the greater the number of previous lists the greater the *proactive* interference. We know this to be true (26) for a formal proactive-inhibition paradigm; it seems a reasonable interpretation for the data of Fig. 3. That there are minor sources of variance still involved I do not deny. Some of the variation can be rationalized, but that is not the purpose of this report. The point I wish to make is the obvious one of the relationship between number of previous lists learned—lists which presumably had no intentionally built-in similarity—and amount of forgetting. If you like to think in correlational terms, the rank-order correlation between the two variables is  $-.91$  for the 14 points of Fig. 3.

It may be of interest to the historian that, of the studies published before 1942 which met the criteria I imposed, I did not find a single one in which sub-

jects had not been given at least one practice task before starting experimental conditions, and in most cases the subjects had several practice lists and several experimental conditions. Gibson's study (1942) was the first I found in which subjects served in only one condition and were not given practice tasks. I think it is apparent that the design proclivities of the 1920s and 1930s have been largely responsible for the exaggerated picture we have had of the rate of forgetting of rote-learned materials. On the basis of studies performed during the 1920s and 1930s, I have given a rough estimate of forgetting as being 75 per cent over 24 hours, recall being 25 per cent. On the basis of modern studies in which the subject has learned no previous lists—where there is no proactive inhibition from previous laboratory tasks—a rough estimate would be that forgetting is 25 per cent; recall is 75 per cent. The values are reversed. (If in the above and subsequent discussion my use of percentage values as if I were dealing with a cardinal or extensive scale is dis-

turbing, I will say only that it makes the picture easier to grasp, and in my opinion no critical distortion results.)

Before taking the next major step, I would like to point out a few other observations which serve to support my general point that proactive inhibition from laboratory tasks has been the major cause of forgetting in the more classical studies. The first illustration I shall give exemplifies the point that when subjects have served in several conditions, forgetting after relatively short periods of time is greater than after 24 hours if the subject has served in only one condition. In the Youtz study to which I have already referred, other conditions were employed in which recall was taken after short intervals. After 20 minutes recall was 74 per cent, about what it is after 24 hours if the subject has not served in a series of conditions. After two hours recall was 32 per cent. In Ward's (34) well-known reminiscence experiment, subjects who on the average had learned ten previous lists showed a recall of only 64 per cent after 20 minutes.

In the famous Jenkins-Dallenbach (13) study on retention following sleep and following waking, two subjects were used. One subject learned a total of 61 lists and the other 62 in addition to several practice lists. Roughly, then, if the order of the conditions was randomized, approximately 30 lists had been learned prior to the learning of a list for a given experimental condition. Recall after eight waking hours for one subject was 4 per cent and for the other 14 per cent. Even after sleeping for eight hours the recall was only 55 per cent and 58 per cent.

I have said that an interpolated list can produce severe forgetting. However, in one study (1), using the A-B, A-C paradigm for original and interpolated learning, but using subjects who had never served in any previous con-

ditions, recall of the original list was 46 per cent after 48 hours, and in another comparable study (24), 42 per cent. Thus, the loss is not nearly as great as in the classical studies I have cited where there was no interpolated learning in the laboratory.

My conclusion at this point is that, in terms of the gross analysis I have made, the amount of forgetting which might be attributed to interference from tasks learned outside the laboratory has been "reduced" from 75 per cent to about 25 per cent. I shall proceed in the next section to see if we have grounds for reducing this estimate still more. In passing on to this section, however, let me say that the study of factors which influence proactive inhibition in these counterbalanced studies is a perfectly legitimate and important area of study. I mention this because in the subsequent discussion I am going to deal only with the case where a subject has learned a single list in the laboratory, and I do not want to leave the impression that we should now and forevermore drop the study of interference produced by previous laboratory tasks. Indeed, as will be seen shortly, it is my opinion that we should increase these studies for the simple reason that the proactive paradigm provides a more realistic one than does the retroactive paradigm.

When the subject learns and recalls a single list in the laboratory, I have given an estimate of 25 per cent as being the amount forgotten over 24 hours. When, as shown above, we calculate percentage forgotten of lists learned to one perfect trial, the assumption is that had the subjects been given an immediate recall trial, the list would have been perfectly recalled. This, of course, is simply not true. The major factor determining how much error is introduced by this criterion-percentage method is probably the difficulty of the task. In general,



the overestimation of forgetting by the percentage method will be directly related to the difficulty of the task. Thus, the more slowly the learning approaches a given criterion, the greater the drop on the trial immediately after the criterion trial. Data from a study by Runquist (24), using eight paired adjectives (a comparatively easy task), shows that amount of forgetting is overestimated by about 10 per cent. In a study (32) using very difficult consonant syllables, the overestimation was approximately 20 per cent. To be conservative, assume that on the average the percentage method of reporting recall overestimates the amount forgotten by 10 per cent. If we subtract this from the 25 per cent assumed above, the forgetting is now re-estimated as being 15 per cent over 24 hours. That is to say, an interference theory, or any other form of theory, has to account for a very small amount of forgetting as compared with the amount traditionally cited.

What are the implications of so greatly "reducing" the amount of forgetting? There are at least three implications which I feel are worth pointing out. First, if one wishes to hold to an interference theory of forgetting (as I do), it seems plausible to assert that this amount of forgetting could be produced from learning which has taken place outside of the laboratory. Furthermore, it seems likely that such interference must result primarily from proactive interference. This seems likely on a simple probability basis. A 20-year-old college student will more likely have learned something during his 20 years prior to coming to the laboratory that will interfere with his retention than he will during the 24 hours between the learning and retention test. However, the longer the retention interval the more important will retroactive in-

terference become relative to proactive interferences.

The second implication is that these data may suggest greater homogeneity or continuity in memorial processes than hitherto supposed. Although no one has adequately solved the measurement problem of how to make comparisons of retention among conditioned responses, prose material, motor tasks, concept learning, and rote-learned tasks, the gross comparisons have indicated that rote-learned tasks were forgotten much more rapidly than these other tasks. But the rote-learning data used for comparison have been those derived with the classical design in which the forgetting over 24 hours is approximately 75 per cent. If we take the revised estimate of 15 per cent, the discrepancies among tasks become considerably less.

The third implication of the revised estimate of rate of forgetting is that the number of variables which appreciably influence rate of forgetting must be sharply limited. While this statement does not inevitably follow from the analyses I have made, the current evidence strongly supports the statement. I want to turn to the final section of this paper which will consist of a review of the influence of some of the variables which are or have been thought to be related to rate of forgetting. In considering these variables, it is well to keep in mind that a variable which produces only a small difference in forgetting is important if one is interested in accounting for the 15 per cent assumed now as the loss over 24 hours. If appropriate for a given variable, I will indicate where it fits into an interference theory, although in no case will I endeavor to handle the details of such a theory.

*Time.* Passage of time between learning and recall is the critical defining variable for forgetting. Manipulation of this variable provides the basic data for

which a theory must account. Previously, our conception of rate of forgetting as a function of time has been tied to the Ebbinghaus curve. If the analysis made earlier is correct, this curve does not give us the basic data we need. In short, we must start all over and derive a retention curve over time when the subjects have learned no previous materials in the laboratory. It is apparent that I expect the fall in this curve over time to be relatively small.

In conjunction with time as an independent variable, we must, in explanations of forgetting, consider why sleep retards the processes responsible for forgetting. My conception, which does not really explain anything, is that since forgetting is largely produced by proactive interference, the amount of time which a subject spends in sleep is simply to be subtracted from the total retention interval when predicting the amount to be forgotten. It is known that proactive interference increases with passage of time (5); sleep, I believe, brings to a standstill whatever these processes are which produce this increase.

*Degree of learning.* We usually say that the better or stronger the learning the more or better the retention. Yet, we do not know whether or not the *rate* of forgetting differs for items of different strength. The experimental problem is a difficult one. What we need is to have a subject learn a single association and measure its decline in strength over time. But this is difficult to carry out with verbal material, since almost of necessity we must have the subject learn a series of associations, to make it a reasonable task. And, when a series of associations are learned, complications arise from interaction effects among associations of different strength. Nevertheless, we may expect, on the basis of evidence from a wide variety of studies, that given a constant degree of similarity, the effective interference varies as

some function of the strength of associations.

*Distribution of practice.* It is a fact that distribution of practice during acquisition influences retention of verbal materials. The facts of the case seem to be as follows. If the subject has not learned previous lists in the laboratory, massed practice gives equal or better retention than does distributed practice. If, on the other hand, the subject has learned a number of previous lists, distributed practice will facilitate retention (32). We do not have the theoretical solution to these facts. The point I wish to make here is that whether or not distribution of learning inhibits or facilitates retention depends upon the amount of interference from previous learning. It is reasonable to expect, therefore, that the solution to the problem will come via principles handling interference in general. I might also say that a theoretical solution to this problem will also provide a solution for Jost's laws.

*Similarity.* Amount of interference from other tasks is closely tied to similarity. This similarity must be conceived of as similarity among materials as such and also situational similarity (4). When we turn to similarity within a task, the situation is not quite so clear. Empirically and theoretically (8) one would expect that intratask similarity would be a very relevant variable in forgetting. As discussed elsewhere (31), however, variation in intratask similarity almost inevitably leads to variations in intertask similarity. We do know from a recent study (33) that with material of low meaningfulness forgetting is significantly greater with high intralist similarity than with low. While the difference in magnitude is only about 8 per cent, when we are trying to account for a total loss of 15 per cent, this amount becomes a major matter.

*Meaningfulness.* The belief has long



been held that the more meaningful the material the better the retention—the less the forgetting. Osgood (21) has pointed out that if this is true it is difficult for an interference theory to handle. So far as I know, the only direct test of the influence of this variable is a recent study in which retention of syllables of 100 per cent association value was compared with that of zero association value (33). There was no difference in the recall of these syllables. Other less precise evidence would support this finding when comparisons are made among syllables, adjectives, and nouns, as plotted in Fig. 3. However, there is some evidence that materials of very low meaningfulness are forgotten more rapidly than nonsense syllables of zero association value. Consonant syllables, both serial (32) and paired associates (unpublished), show about 50 per cent loss over 24 hours. The study using serial lists was the one mentioned earlier as knowingly omitted from Fig. 3. These syllables, being extremely difficult to learn, allow a correction of about 20 per cent due to criterion overestimation, but even with this much correction the forgetting (30 per cent) is still appreciably more than the estimate we have made for other materials. To invoke the interference theory to account for this discrepancy means that we must demonstrate how interference from other activities could be greater for these consonant syllables than for nonsense syllables, nouns, adjectives, and other materials. Our best guess at the present time is that the sequences of letters in consonant syllables are contrary to other well-established language habits. That is to say, letter sequences which commonly occur in our language are largely different from those in consonant syllables. As a consequence, not only are these consonant syllables very difficult to learn, but forgetting is accelerated by proactive interference from previously

well-learned letter sequences. If subsequent research cannot demonstrate such a source of interference, or if some other source is not specified, an interference theory for this case will be in some trouble.

*Affectivity.* Another task dimension which has received extensive attention is the affective tone of the material. I would also include here the studies attaching unpleasant experiences to some items experimentally and not to others, and measuring retention of these two sets of items. Freud is to a large extent responsible for these studies, but he cannot be held responsible for the malformed methodology which characterizes so many of them. What can one say by way of summarizing these studies? The only conclusion that I can reach is a statistical one, namely, that the occasional positive result found among the scores of studies is about as frequent as one would expect by sampling error, using the 5 per cent level of confidence. Until a reliable body of facts is established for this variable and associated variables, no theoretical evaluation is possible.

*Other variables.* As I indicated earlier, I will not make an exhaustive survey of the variables which may influence rate of forgetting. I have limited myself to variables which have been rather extensively investigated, which have immediate relevance to the interference theory, or for which reliable relationships are available. Nevertheless, I would like to mention briefly some of these other variables. There is the matter of *warm-up* before recall; some investigators find that this reduces forgetting (12); others, under as nearly replicated conditions as is possible to obtain, do not (23). Some resolution must be found for these flat contradictions. It seems perfectly reasonable, however, that inadequate set or context differences could reduce recall. Indeed, an



interference theory would predict this forgetting if the set or context stimuli are appreciably different from those prevailing at the time of learning. In our laboratory we try to reinstate the learning set by careful instructions, and we simply do not find decrements that might be attributed to inadequate set. For example, in a recent study (33) subjects were given a 24-hour recall of a serial list after learning to one perfect trial. I think we would expect that the first item in the list would suffer the greatest decrement due to inadequate set, yet this item showed only .7 per cent loss. But let it be clear that when we are attempting to account for the 15 per cent loss over 24 hours, we should not overlook any possible source for this loss.

Thus far I have not said anything about forgetting as a function of characteristics of the subject, that is, the personality or intellectual characteristics. As far as I have been able to determine, there is not a single valid study which shows that such variables have an appreciable influence on forgetting. Many studies have shown differences in learning as a function of these variables, but not differences in rate of forgetting. Surely there must be some such variables. We do know that if subjects are severely insulted, made to feel stupid, or generally led to believe that they have no justification for continued existence on the earth just before they are asked to recall, they will show losses (e.g., 25, 38), but even the influence of this kind of psychological beating is short lived. Somehow I have never felt that such findings need explanation by a theory used to explain the other facts of forgetting.

Concerning the causes of forgetting, let me sum up in a somewhat more dogmatic fashion than is probably justified. One of the assumptions of science is finite causality. Everything cannot in-

fluence everything else. To me, the most important implication of the work on forgetting during the last ten years is that this work has markedly *reduced* the number of variables related to forgetting. Correspondingly, I think the theoretical problem has become simpler. It is my belief that we can narrow down the cause of forgetting to interference from previously learned habits, from habits being currently learned, and from habits we have yet to learn. The amount of this interference is primarily a function of similarity and associative strength, the latter being important because it interacts with similarity.

#### SUMMARY

This paper deals with issues in the forgetting of rote-learned materials. An analysis of the current evidence suggests that the classical Ebbinghaus curve of forgetting is primarily a function of interference from materials learned previously in the laboratory. When this source of interference is removed, forgetting decreases from about 75 per cent over 24 hours to about 25 per cent. This latter figure can be reduced by at least 10 per cent by other methodological considerations, leaving 15 per cent as an estimate of the forgetting over 24 hours. This estimate will vary somewhat as a function of intratask similarity, distributed practice, and with very low meaningful material. But the overall evidence suggests that similarity with other material and situational similarity are by far the most critical factors in forgetting. Such evidence is consonant with a general interference theory, although the details of such a theory were not presented here.

#### REFERENCES

1. ARCHER, E. J., & UNDERWOOD, B. J. Retroactive inhibition of verbal associations as a multiple function of temporal point of interpolation and degree of interpo-

- lated learning. *J. exp. Psychol.*, 1951, 42, 283-290.
2. BARTLETT, F. C. *Remembering: a study in experimental and social psychology*. London: Cambridge Univer. Press, 1932.
3. BELMONT, L., & BIRCH, H. G. Re-individualizing the repression hypothesis. *J. abnorm. soc. Psychol.*, 1951, 46, 226-235.
4. BILODEAU, I. McD., & SCHLOSBERG, H. Similarity in stimulating conditions as a variable in retroactive inhibition. *J. exp. Psychol.*, 1951, 41, 199-204.
5. BRIGGS, G. E. Acquisition, extinction, and recovery functions in retroactive inhibition. *J. exp. Psychol.*, 1954, 47, 285-293.
6. CHENG, N. Y. Retroactive effect and degree of similarity. *J. exp. Psychol.*, 1929, 12, 444-458.
7. EBBINGHAUS, H. *Memory: a contribution to experimental psychology*. (Trans. by H. A. Ruger, and C. E. Bussenius) New York: Bureau of Publications, Teachers College, Columbia Univer., 1913.
8. GIBSON, ELEANOR J. A systematic application of the concepts of generalization and differentiation to verbal learning. *Psychol. Rev.*, 1940, 47, 196-229.
9. GIBSON, ELEANOR J. Intra-list generalization as a factor in verbal learning. *J. exp. Psychol.*, 1942, 30, 185-200.
10. GREENBERG, R., & UNDERWOOD, B. J. Retention as a function of stage of practice. *J. exp. Psychol.*, 1950, 40, 452-457.
11. HOVLAND, C. I. Experimental studies in rote-learning theory. VI. Comparison of retention following learning to same criterion by massed and distributed practice. *J. exp. Psychol.*, 1940, 26, 568-587.
12. IRON, A. L. The relation of "set" to retention. *Psychol. Rev.*, 1948, 55, 336-341.
13. JENKINS, J. G., & DALLENBACH, K. M. Oblivescence during sleep and waking. *Amer. J. Psychol.*, 1924, 35, 605-612.
14. JOHNSON, L. M. The relative effect of a time interval upon learning and retention. *J. exp. Psychol.*, 1939, 24, 169-179.
15. KATONA, G. *Organizing and memorizing: studies in the psychology of learning and teaching*. New York: Columbia Univer. Press, 1940.
16. KRUEGER, W. C. F. The effect of over-learning on retention. *J. exp. Psychol.*, 1929, 12, 71-78.
17. LESTER, O. P. Mental set in relation to retroactive inhibition. *J. exp. Psychol.*, 1932, 15, 681-699.
18. LUH, C. W. The conditions of retention. *Psychol. Monogr.*, 1922, 31, No. 3 (Whole No. 142).
19. MCGEOCH, J. A. Forgetting and the law of disuse. *Psychol. Rev.*, 1932, 39, 352-370.
20. MCGEOCH, J. A. *The psychology of human learning*. New York: Longmans, Green, 1942.
21. OSGOOD, C. E. *Method and theory in experimental psychology*. New York: Oxford Univer. Press, 1953.
22. RAPAPORT, D. Emotions and memory. *Psychol. Rev.*, 1943, 50, 234-243.
23. ROCKWAY, M. R., & DUNCAN, C. P. Pre-recall warming-up in verbal retention. *J. exp. Psychol.*, 1952, 43, 305-312.
24. RUNQUIST, W. Retention of verbal associations as a function of interference and strength. Unpublished doctor's dissertation, Northwestern Univer., 1956.
25. RUSSELL, W. A. Retention of verbal material as a function of motivating instructions and experimentally-induced failure. *J. exp. Psychol.*, 1952, 43, 207-216.
26. UNDERWOOD, B. J. The effect of successive interpolations on retroactive and proactive inhibition. *Psychol. Monogr.*, 1945, 59, No. 3 (Whole No. 273).
27. UNDERWOOD, B. J. Studies of distributed practice: VII. Learning and retention of serial nonsense lists as a function of intralist similarity. *J. exp. Psychol.*, 1952, 44, 80-87.
28. UNDERWOOD, B. J. Studies of distributed practice: VIII. Learning and retention of paired nonsense syllables as a function of intralist similarity. *J. exp. Psychol.*, 1953, 45, 133-142.
29. UNDERWOOD, B. J. Studies of distributed practice: IX. Learning and retention of paired adjectives as a function of intralist similarity. *J. exp. Psychol.*, 1953, 45, 143-149.
30. UNDERWOOD, B. J. Studies of distributed practice: X. The influence of intralist similarity on learning and retention of serial adjective lists. *J. exp. Psychol.*, 1953, 45, 253-259.
31. UNDERWOOD, R. J. Intralist similarity in verbal learning and retention. *Psychol. Rev.*, 1954, 3, 160-166.
32. UNDERWOOD, B. J., & RICHARDSON, J. Studies of distributed practice: XIII. Interlist interference and the retention of serial nonsense lists. *J. exp. Psychol.*, 1955, 50, 39-46.

33. UNDERWOOD, B. J., & RICHARDSON, J. The influence of meaningfulness, intralist similarity, and serial position on retention. *J. exp. Psychol.*, 1956, 52, 119-126.
34. WARD, L. B. Reminiscence and rote learning. *Psychol. Monogr.*, 1937, 49, No. 4 (Whole No. 220).
35. WEISS, W., & MARGOLIUS, G. The effect of context stimuli on learning and retention. *J. exp. Psychol.*, 1954, 48, 318-322.
36. WILLIAMS, M. The effects of experimentally induced needs upon retention. *J. exp. Psychol.*, 1950, 40, 139-151.
37. YOUTZ, ADELLA C. An experimental evaluation of Jost's laws. *Psychol. Monogr.*, 1941, 53, No. 1 (Whole No. 238).
38. ZELLER, A. F. An experimental analogue of repression: III. The effect of induced failure and success on memory measured by recall. *J. exp. Psychol.*, 1951, 42, 32-38.

(Received April 23, 1956)



## A CRITIQUE OF KÖHLER'S THEORY OF ASSOCIATION

LEO POSTMAN AND DONALD A. RILEY

*University of California*

In 1941 Köhler published a paper entitled, "On the nature of associations" (3). After restating the traditional opposition between *association* and *organization*, Köhler presented what he considered to be decisive experimental evidence in favor of a Gestalt theory of learning. Köhler's findings do, indeed, appear to raise serious difficulties for association theory. We shall try to show that these difficulties are only apparent, and that the experimental facts can be fully understood in terms of principles of associative learning.<sup>1</sup>

*Köhler's theory of association.* In his critical analysis of the process of association Köhler focuses on the "relation which the characteristics of one item bear to the characteristics of the other" (3, p. 490). According to Köhler, this relation is a matter of indifference for association theory. The formation of connections is a function of contiguity, regardless of the nature of the items. For Gestalt theory, on the other hand, the relation between the items is a crucial determinant of both learning and retention. The relation between the items is critical, since learning is considered to be a matter of perceptual organization. Perceptual organizations ("primary experiences"), in turn, are preserved in the nervous system as memory traces. If the perceptual experience is "unitary" by virtue of the proximity or similarity of the compo-

nent parts, the corresponding memory trace has the same "unitary" character.

Given these assumptions, the fact of association by contiguity can be reinterpreted as a special case of organization. Suppose a unitary trace has been formed, and part of such a trace is reactivated at a later time. "Because of the unitary character of the trace, this excitation will spread more easily within the trace than to other regions of the tissue" (3, p. 493). Association, Köhler concluded, is "simply coherence within the unitary trace of a unitary experience." Association by contiguity is, therefore, a secondary concept subsidiary to the more general concept of organization.

*Experimental evidence for Köhler's theory.* This argument has a clear-cut empirical implication. Any principle of perceptual organization is necessarily also a principle of association. In his experiments Köhler chose the variable of *similarity* to demonstrate the validity of this implication. Similarity favors perceptual grouping, i.e., similar items tend to form unitary configurations. By the same token, similarity of cue and response should favor the formation of connections in paired-associate learning. What appears to be positive evidence for the hypothesis was obtained in a series of related experiments.

In Köhler's first study, two series of six pairs each were used: (a) a series of *homogeneous* pairs consisting of two pairs of nonsense syllables, two pairs of two-place numbers, and two pairs of nonsense figures; (b) a series of *heterogeneous* pairs containing one of each of the six possible combinations of the three types of items. Two presentations

<sup>1</sup> Professor Köhler reaffirmed the crucial significance which he attaches to his findings during his Hitchcock Lectures, delivered at the University of California, Berkeley, in the spring of 1955. It was because these studies were again offered as decisive proof for his theory of association that our interest in this problem was aroused.

of the series (apparently in a constant serial order for a given *S*) were followed by one test trial using the method of paired associates. All *Ss* learned both series, with the order of the two series counterbalanced. A substantially higher percentage of homogeneous than heterogeneous items was recalled. The second experiment introduced two modifications. First, the response items in the two series were identical, which they had not been in the first study. Second, independent groups learned the homogeneous and heterogeneous series. Again, homogeneous items were recalled much better than heterogeneous items. The third experiment was designed to show that the critical variable was the similarity between the members of *individual pairs* rather than the composition of the total series. Homogeneous and heterogeneous pairs were combined in one series. As before, recall was considerably higher for homogeneous than heterogeneous items. In the final experiment, homogeneity and heterogeneity were defined in terms of the physical resemblance of the members of a pair. In the heterogeneous series, the cue and response members differed with respect to size and color; in the homogeneous series there were no such differences in physical appearance. Such "extrinsic" similarity of the items again favored recall, although the differences were smaller than in the earlier experiments in which the "intrinsic" similarity of the items was varied.

*Did Köhler test his own hypothesis?* Köhler's theoretical analysis, his derivation of the empirical predictions, and his experimental treatment are all open to serious questions. The central point of Köhler's position is the assumption of an isomorphic relation between perceptual organization and memory trace. Throughout his analysis, Köhler equates the arousal of memory traces with the spread of excitation through "ganglionic

tissue." There is no doubt, therefore, that a *spatial* correspondence between percept and memory trace is assumed. Serious difficulties arise in the application of this model of cortical projection to the dimensions of similarity manipulated in the experiments.

In the first three experiments, cue-response pairs were considered "similar" or "homogeneous" if both members of a pair belonged to the same meaning class, i.e., syllables, numbers, or geometric figures. Such classifications are arbitrary conventions. As *geometric patterns*, there are numbers which are more similar to some letters than they are to other numbers. Consider, for example, the number 0 (zero) and the letter *O*; or the number 1 (one) and the letter *l*. Clearly, Köhler chose to define perceptual similarity in terms of the *arbitrary, learned* categories to which the individual items can be assigned. Such a conception of similarity is not very different from Thorndike's principle of "belongingness," which Köhler rejects as inadequate because it refers to arbitrary rather than intrinsic relationships among the stimuli. In practice, similarity for Köhler means, "this goes with that," just as it did for Thorndike. In addition, Köhler is forced to assume that the cortical traces of items belonging to the same conventional class interact more readily in the "ganglionic tissue" than do the traces of physically similar items belonging to different arbitrary classes! Such neurological speculations are highly implausible. We conclude that similarity defined in terms of class membership has no relevance to physiological hypotheses derived from the principle of isomorphism. In using such a definition of similarity, Köhler failed to test his own hypothesis.

Color and size, which were used to define similarity in the fourth experiment, at first glance seem to be more



plausible as determinants of cortical interaction than is membership in an arbitrary meaning class. Size at least is a physical characteristic to which the hypothesis of (spatial) isomorphism is directly relevant. There is no clear evidence for the cortical localization of color (4). What type of spatial differentiation among traces is to be assumed when stimuli vary in both size and color, as they did in Köhler's experiment? There is no answer to this question based on physiological evidence, and Köhler's theory is inevitably silent on this point. At best, then, the results of the experiment on physical similarity may provide indirect support for a vague physiological hypothesis. Since this experiment is the only one that may have some relevance to the theory, its reproducibility assumes critical importance.

*Köhler's analysis of association theory.* Granted that Köhler largely failed to test his own hypothesis, what are the implications of his findings for association theory? The phrase "association theory" covers a multitude of theoretical positions, but Köhler focuses his criticism on two points which he apparently considers common to all associationistic approaches. First, the relationship between the stimulus and response items is a matter of indifference. Contiguity is the basic determinant of association. Second, Köhler suggests that association theorists agree on a *physiological theory* which is coordinate with the conception of association as an "indifferent bond." According to Köhler, association theorists identify learning with changes in conductivity along "individual nerve fibers" linking the central representations of S and R. In the light of his experimental findings, Köhler finds such a neurological theory grossly inadequate to account for the facts of learning.

This physiological theory of associa-

tion represents a straw man of the critic's own construction, or, at best, a confusion of outdated historical doctrine with current positions. It is true that Thorndike (e.g., 9) identified S-R connections with neural bonds, but such physiological speculations remained incidental to Thorndike's main theoretical effort and did not exert any direct influence on his experimental work. Beyond locating the mechanisms of association in the "life processes" of the neurons, Thorndike never offered detailed physiological explanations of specific experimental results.

In any event, contemporary association theory is certainly not built upon Thorndikian neurology; in fact, it is usually not built on any neurology at all.<sup>2</sup> The major constructs of modern learning theories, such as habit or expectancy, are symbolic constructs without any specific physiological connotation. Both Hull and Guthrie have been prone to describe stimulus effects and responses in "quasi-physiological language," but have no specific physiological theory of association. Other association theorists bypass the problem of physiological translation even more completely. There is no speculation concerning physiological processes in Tolman or Skinner. Indeed, Skinner explicitly rejects any recourse to neurological hypotheses as premature and misleading (8, p. 193 f.). Certainly "functionalists" such as McGeoch never veered from the path of strictly behavioral analysis (6). We do not know what association theorist was the target of Köhler's attack. The Thorndike of *Animal Intelligence* (1898)? The Thorndike of *Human Learning* (1931)?

<sup>2</sup> The work of Hebb (2) is an important exception. His theory emphasizes complex interactions in the nervous system, and bears little resemblance to the type of physiological hypotheses that Köhler chose to ascribe to associationists.



Hull? Tolman? Skinner? McGeoch? Guthrie? The criticism is either out of date or does not apply. We conclude that Köhler's experiments have no bearing on the physiological hypotheses of associationism. It remains for us to examine the results in relation to associationist behavior theory.

*Perceptual similarity or generalization?* Köhler suggests that there are no principles of association which can account for his experimental results. The difficulty arises, however, from Köhler's formulation of the problem rather than from an inadequacy of associationistic principles of learning. Köhler chose to analyze the learning of a *series* of paired associates with exclusive reference to the properties of *individual* pairs. He makes this point quite explicitly when he says, "... our problem refers to *individual* pairs and the degree of resemblance between their members" (3, p. 498). When the learning material consists of a set of items, any analysis which fails, as Köhler's does, to take account of the characteristics of the set will be necessarily incomplete. The critical importance of intralist relationships has long been recognized. Gibson (1) has systematized the role of intralist effects by an application of the principles of generalization and differentiation. Gibson's thesis is that the course of learning is a function of the degree of generalization among the items to be learned. Generalization is, in turn, a function of similarity. To the extent that generalization favors incorrect responses, i.e., when similar items require different responses, learning is delayed. To the extent that generalization favors correct responses, i.e., when similar items require the same response, learning is facilitated. It will be our contention that the advantage of homogeneous over heterogeneous pairs is a function of intralist generalization.

Let us turn first to Köhler's two ex-

periments in which Ss learned separate series of homogeneous and heterogeneous pairs composed of nonsense syllables, numbers, and geometric designs. We must consider not only the similarity between the two members of each pair but also the similarity among the cue items and among the response items. Both the cue items and the response items will be categorized by S as belonging to three classes: syllables, numbers, and designs. This grouping will be mediated by the differential responses which S is assumed to make to the individual items (5). These differential responses include labeling or identification of the items as members of one of the three classes. By virtue of a common differential response, items belonging to the same class will be more similar to each other than they will be to members of other classes.

Mediated stimulus generalization will occur primarily among similar items, i.e., items belonging to the same class. Thus, there will be more generalization between two syllable cues than there will be between a syllable cue and a number or design cue. When the pairs are homogeneous, similar cues will all require responses belonging to *one* class, e.g., all syllable cues will require syllable responses. Generalization will strengthen S's tendency always to give a response belonging to the appropriate class. He then merely needs to differentiate between the items *within* a class. By contrast, when the pairs are heterogeneous, similar cues require responses from *two* different classes. Generalization between cues will lead to responses from the incorrect class, e.g., responding to a syllable with a design when a number is correct, etc. In order to learn the series of heterogeneous pairs, S must eliminate two types of errors: (a) giving responses from the wrong class, and (b) confusion between the items within a given class. The S learning a homo-

geneous series has to overcome only the second type of error. Hence, homogeneous pairs will be learned faster than heterogeneous pairs. The effects of generalization may also be described in somewhat different terms. When the series is homogeneous, generalization tendencies conform to an unequivocal *rule of response selection* (syllables go with syllables, numbers with numbers and designs with designs). When the series is heterogeneous, the rule is equivocal. Responses may not come from the same class as the cue items, but they may come from one of two other classes. The more equivocal the rule, the more opportunity there is for intralist errors.

Our argument may be supplemented by a more formal analysis of the associative patterns established during the acquisition of homogeneous and heterogeneous lists. Refer to Figure 1 (left) which presents the assumed pattern of associations for Köhler's list of homogeneous pairs. A differential response ( $r$ ) is linked with each cue item ( $S$ ). The performance of this differential response results in response-produced stimulation ( $s$ ). Differential responses and response-produced stimuli are more sim-

ilar for members of the same class of items than they are for members of different classes. Response items are similarly differentiated in terms of  $r$ - $s$  patterns. In the course of learning, two types of associations are established: (a) "direct" associations between  $S$  and  $R$ , and (b) associations between  $S$  and  $R$  which are mediated by the differential response to cue items and response items, e.g.,  $S_{x_1} \rightarrow r_{x_1} \rightarrow s_{x_1} \rightarrow r_{x_2} \rightarrow s_{x_2} \rightarrow R_{x_2}$ , where  $x_1$  and  $x_2$  are, say, two syllables forming a homogeneous pair. Generalized responses arise by virtue of similarities among the  $S$ s as well as among the  $s$ s. The dashed arrows in Figure 1 indicate examples of generalized responses mediated by  $r$ - $s$  sequences. Note also that the similarity between  $r$ - $s$  sequences connected with *response items within a given class* will result in response generalization. It is clear that the generalized responses conform to a rule of response selection, and tend to restrict intralist errors to confusion of items within a class. Figure 1 (right) presents the pattern of associations for the heterogeneous list. Applying the same analysis as before, we find that generalization fails to follow an un-

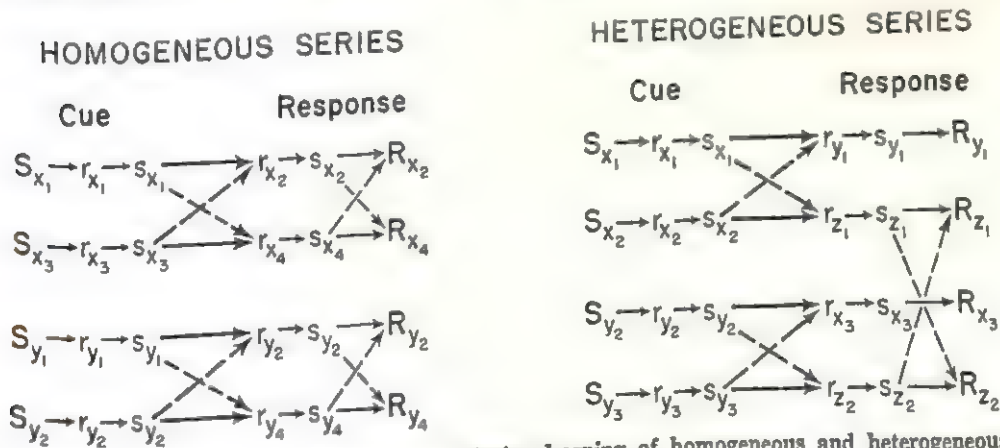


FIG. 1. Mediated associations formed during learning of homogeneous and heterogeneous lists.  $S$  = cue;  $r$  = differential response;  $s$  = response-produced stimulation resulting from  $r$ ;  $R$  = overt verbal response. The subscripts,  $x$ ,  $y$  and  $z$ , denote classes of items, i.e., syllables, numbers, and responses. Dashed arrows denote generalization tendencies. Only mediated associations are shown, but direct  $S$ - $R$  connections are also assumed to be formed. To simplify exposition, only four of the six pairs in Köhler's series are represented.



equivocal rule of response selection and maximizes the opportunity for intralist errors, i.e., leads to errors both between classes and within classes.

Finally, we must consider the effects of transfer of training, i.e., the influence of the response dispositions with which *S* enters into the experimental situation. Köhler anticipates that association theorists would be inclined to argue that "syllables go with other syllables in speech, writing and reading, and that, similarly, numbers go with numbers in daily experience" (3, p. 499).<sup>3</sup> We agree. Köhler dismisses the importance of this fact primarily because "previous associations of this kind will constitute vague reproduction tendencies in many different directions" (3, p. 499). Hence, there could be no positive transfer to the acquisition of *specific* homogeneous pairs; indeed, the transfer may be negative. Here we disagree. We suggest that *S*'s previous experiences with sequences of syllables, numbers, etc., will result in a significant bias for a *rule of response selection*, viz., the rule that like goes with like. In terms of the symbols of Figure 1, what will be significant is not the transfer of specific S-R linkages but rather linkages of the type,  $s_w \rightarrow r_w$ . There will be a generalized disposition to respond to syllables with syllables, to numbers with numbers, etc. When the series consists of homogeneous pairs, this disposition conforms to the rule of the series, and will favor correct performance. When the series consists of heterogeneous pairs, this disposition will delay the acquisition of correct responses. Thus, transfer of training will serve to intensify the differences produced by intralist effects and increase

the advantage of the homogeneous series.

Let us now consider Köhler's third experiment, in which homogeneous and heterogeneous pairs appeared in the *same* series. Only two types of items—syllables and designs—were used. There were four pairs: two homogeneous and two heterogeneous. The series does not obey a rule of response selection, i.e., intralist generalization does not favor the homogeneous pairs. The assumption of an initial response bias does, however, lead to the prediction that homogeneous pairs will be learned faster than heterogeneous pairs. In the case of homogeneous pairs, *S* can continue to apply the rule, and needs only to differentiate between the specific items in a given class. The heterogeneous pairs must, on the other hand, be learned in spite of the rule of response selection which *S* favors.

*Experimental implications of the associationistic interpretation.* So far, our reinterpretation of Köhler's results may be considered *ad hoc*. Fortunately, it is possible to decide between his interpretation and ours by experimental tests. We have asserted that in Köhler's experiments the operation of a rule of response selection favored homogeneous pairs over heterogeneous pairs. Köhler's interpretation hinges on the similarity between the members of *individual* pairs. To decide between these alternatives, it is necessary to compare the acquisition of a homogeneous and heterogeneous series when *both* obey a rule of response selection. This will be true when the series are composed of two rather than three classes of items. For the homogeneous series, the rule of response selection, then, is that like items always go together; for the heterogeneous series, unlike items always go together. In both cases the class membership of the stimulus fully predicts the class membership of the response. Ac-

<sup>3</sup> Köhler also suggests that no transfer could be assumed in the case of geometric designs. The reason for this statement is not clear. Abstract designs may be less frequent than syllables and numbers but are apt to be encountered in groups or sequences.



cording to Köhler's hypothesis, the difference in favor of homogeneous pairs should occur as clearly as in the earlier experiments. According to our hypothesis, the difference should be eliminated, except for the effects of initial response bias. The differential effects of initial bias should, however, be transitory, since the rule that like items go together would never lead to correct responses in the learning of the heterogeneous series. In addition, our analysis predicts, and Köhler's does not, that *Ss'* errors will reflect the rule of response selection. Our prediction follows directly from the generalization hypothesis. Köhler's interpretation does not provide for systematic variation in intralist errors as a function of the characteristics of the series, since his argument makes reference only to the composition of individual pairs. Our hypothesis implies, on the other hand, that the predominant type of error will vary directly with the rule of response selection.

In the case of a series composed of both homogeneous and heterogeneous pairs, the assumption of an initial response bias leads to the prediction that the homogeneous pairs will be learned faster. Here our prediction agrees with Köhler's. Again, however, our analysis can, and Köhler's cannot, generate a specific prediction concerning the distribution of errors. Since *S* has an initial disposition to link like items together, the errors at the beginning of practice should be predominantly in accordance with this rule. As training continues, and heterogeneous pairs begin to be learned, the relative frequency of such errors should decrease steadily until the total list has been learned.

Our reinterpretation of Köhler's results does not apply to his final experiment in which the physical resemblance of the members of the pairs was varied. As indicated earlier, the results of that experiment were much less clear-cut

than those of the other studies. We doubted the reproducibility of these results, and, as we shall see, our suspicions were confirmed.

Since our main concern is with the theoretical issues, we have refrained from a detailed criticism of Köhler's experimental method. Some of the most serious difficulties must, however, be mentioned briefly. (a) A fixed number of three presentations was used. Since difficulty varied from series to series, it is probable that the performance of the various groups was measured at different stages of learning. (b) For a given *S*, the order of presentation appears to have been constant from trial to trial. This is a highly unusual procedure for an investigation of paired-associate learning. A fixed order of presentation leads to serial connections among responses. It then becomes impossible to say to what extent performance depends on cue-response linkages, and to what extent on response-response linkages. (c) The analysis fails to take account of *Ss'* errors which offer important clues to the nature of the associations formed during training. The questions raised by Köhler's analysis can be answered adequately only by charting the systematic changes in *S*'s behavior throughout the course of learning to a criterion. In short, the experimental problem calls for use of the standard procedure of paired-associate learning (7, p. 15).

Our own experiments were, therefore, designed to answer the following questions under standard conditions of paired-associate learning: (a) Do series of homogeneous pairs and series of heterogeneous pairs differ in speed of acquisition when both series obey a rule of response selection? (b) Is there a difference in speed of acquisition between homogeneous and heterogeneous pairs within a single series? In both (a) and (b) the terms "homogeneous" and "heterogeneous" refer to the clas-

sification of items by meaning. (c) Are Köhler's findings concerning the effects of the physical resemblance of items within pairs reproducible? (d) What light is thrown on the relative merits of Köhler's theory and an associationistic interpretation by a detailed analysis of the course of acquisition and the distribution of the errors?

*Experimental tests.* We used the procedure customary in the investigation of paired-associate learning. The cue-response pairs were presented on a Hull-type memory drum. The cue item was presented for two seconds, followed immediately by a two-second exposure of the complete pair. There was an interval of eight seconds between trials.<sup>4</sup> After an initial reading of the list, the anticipation method was used to a criterion of one perfect recitation. Four different orders of the pairs were used in rotation.

In Experiment I we compared learning of a homogeneous list and a heterogeneous one when both lists obey a rule of response selection. Therefore, only two types of items—syllables and numbers—were used. The pool of learning items consisted of eight nonsense syllables of 6.67 per cent to 40.00 per cent association value, and eight two-digit numbers. The homogeneous series consisted of four pairs of syllables and four pairs of numbers. The heterogeneous series consisted of four syllable-number pairs and four number-syllable pairs. Both the cue and response items were identical for the two series; the only difference between the homogeneous and heterogeneous lists was in the way the items were combined. Thirty-two Ss were used, 16 under each of the two conditions.

The following categories were used in

the analysis of Ss' performance: (a) correct anticipations, (b) failures to respond, and (c) overt errors. Overt errors included intralist intrusions, responses from outside the list, and incomplete responses. All overt errors were divided into two categories: *like errors* and *unlike errors*. A like error occurred when the incorrect response belonged to the same class of items as the cue, i.e., an incorrect syllable response to a syllable cue or an incorrect number response to a number cue. An unlike error occurred when the incorrect response belonged to the class of items different from the cue, i.e., an incorrect syllable response to a number cue or an incorrect number response to a syllable cue.

When the lists both conform to a rule of response selection, they are learned with equal speed. The mean number of trials to criterion was 23.88 for the homogeneous list, and 23.00 for the heterogeneous list. The slight superiority of the heterogeneous list (1) is, however, not significant. For *both* lists, Ss' responses conformed to the rule of response selection. Of a total of 807 overt errors made by Ss learning the homogeneous list, only three were unlike errors. Similarly, of the 578 errors made by Ss learning the heterogeneous list, only eight were like errors. The two rules of response selection had, however, different effects on the course of learning.

Figure 2 shows Vincent curves for correct responses, failures to respond, and overt errors under the two conditions.<sup>5</sup> Under the homogeneous condition there are a larger number of correct responses *and* a larger number of

<sup>4</sup> Köhler does not report the temporal intervals used in his experiments, so that we do not know whether our intervals are comparable to his.

<sup>5</sup> For purposes of constructing Vincent curves, the total number of trials for each S was divided into five equal parts. The number of correct responses, failures, and overt errors per trial was then determined for each fifth. The means of these values are plotted in the figure.



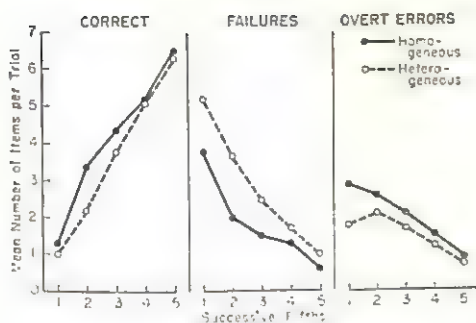


FIG. 2. Vincent curves for correct responses, failures to respond and overt errors in Experiment I.

overt errors than under the heterogeneous condition. Consequently, there are also fewer failures to respond under the homogeneous condition. For each response classification, the mean difference between the homogeneous and heterogeneous condition is significant at either the .05 or the .01 level. The two groups necessarily converge in all respects as the criterion of perfect recitation is approached.

If we were to measure correct responses at some arbitrary point early in learning, we would appear to obtain results similar to Köhler's, i.e., better recall for the homogeneous pairs. It is now clear that to ascribe these results to the facilitating effect of perceptual organization on recall would be a mistake, since Köhler's analysis cannot account for the difference in errors between the two conditions. The higher frequency of both correct and incorrect responses under the homogeneous condition readily falls into place on the assumption that Ss initially favor the rule of response selection appropriate to that condition.

Experiment II compares the learning of homogeneous and heterogeneous pairs within a single series. The same cue and response items were used as in Experiment I, except that the items were so arranged as to form four homogeneous and four heterogeneous pairs. Half the

homogeneous pairs consisted of syllables, the other half consisted of numbers. Similarly, half the heterogeneous pairs consisted of syllable-number combinations, and the other half of number-syllable combinations. Sixteen Ss learned the list under the same conditions as used in Experiment I.

The mean number of trials to criterion was 23.81. The over-all speed of learning was, therefore, almost identical with that in Experiment I. Homogeneous pairs were, however, learned much more rapidly than heterogeneous pairs. The mean number of trials required for correct anticipation of all homogeneous pairs in a single trial was 15.75. In the case of heterogeneous pairs, this criterion was reached after 22.81 trials. This difference is significant ( $p < .01$ ). Note that the number of trials to criterion for homogeneous and heterogeneous pairs is virtually identical when the two types of items appear in separate lists. When both appear in the same list, homogeneous pairs have a striking advantage.

We have predicted speedier learning for homogeneous pairs on the assumption of an initial bias in favor of like responses. The existence of such a bias is clearly demonstrated in Table 1, which lists the mean number of like and unlike errors per trial for successive

TABLE 1  
AVERAGE NUMBER OF LIKE AND UNLIKE ERRORS  
PER TRIAL FOR SUCCESSIVE FIFTHS OF  
LEARNING PERIOD IN EXPERIMENT II

Successive Fifths	Homogeneous Pairs		Heterogeneous Pairs	
	Like Errors	Unlike Errors	Like Errors	Unlike Errors
1	.90	.27	.78	.46
2	.71	.18	1.06	.48
3	.43	.10	.72	.28
4	.32	.12	.54	.58
5	.19	.07	.30	.25



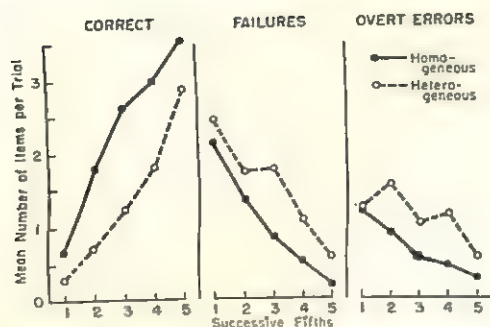


FIG. 3. Vincent curves for correct responses, failures to respond and overt errors in Experiment II.

fifths of the learning period. Like errors predominate, especially in the early stages of learning. While like errors decline steadily over successive fifths, unlike errors show only irregular variations. Progress toward the criterion is signalled primarily by the elimination of the initially frequent like errors. Both the mean difference between like and unlike errors and the interaction between type of error and stage of learning are significant ( $p < .01$ ). We should add that, throughout the course of learning, *both* types of errors are significantly more frequent for the heterogeneous pairs than for the homogeneous pairs.

Figure 3 shows Vincent curves for the different classes of responses in Experiment II. As in Experiment I, we find a larger number of correct responses for homogeneous pairs. In contrast with Experiment I, we now find a greater frequency of overt errors as well as failures for heterogeneous pairs. For each response classification, the mean difference between homogeneous and heterogeneous pairs is significant ( $p < .01$ ). As noted above, like errors predominate for both types of pairs. Such errors continue to be given to the heterogeneous pairs even after the homogeneous pairs have reached a high level of learning. In short, the response bias

favors the learning of homogeneous pairs but produces a sustained high level of errors and slow learning for the heterogeneous pairs.

As predicted, the difference in favor of the homogeneous pairs is considerably enhanced when both types of items appear in the same list. The difference between the two experiments is easily explained on the assumption of an initial response bias. Köhler's analysis could not have predicted the difference between Experiments I and II. In fact, his method of measurement—determination of correct responses regardless of errors after an arbitrary number of trials—revealed no comparable differences in his own study.

In Experiment III we attempted to reproduce Köhler's finding that pairs homogeneous in size and color are learned faster than pairs which are heterogeneous in these respects. The list of cue-response pairs and the conditions of practice were exactly the same as in Experiment II. For one group of 16 Ss, the cue and response items were homogeneous in size and color. For half the group, all items were large ( $\frac{5}{16}$  in. high) and black; for the other half, all items were small ( $\frac{2}{16}$  in. high) and red. For another group of 16 Ss, the cue and response items were heterogeneous in size and color. For half the group, the cue items were large and black, and the response items were small and red; for the other half, the cue items were small and red and the response items were large and black. In the present experiment, therefore, homogeneity and heterogeneity refer to the size and color of the items rather than to their membership in conventional meaning classes.

The mean number of trials to criterion was 26.75 for the homogeneous group, and 23.56 for the heterogeneous group. The difference in favor of the heterogeneous (!) group was not significant. Since Ss were not required to

recall the size and color of the response items, an analysis in terms of like and unlike errors is impossible. Figure 4 shows Vincent curves for the homogeneous and heterogeneous conditions. The slight differences between the two groups in the number of correct responses and overt errors do not approach statistical significance. The mean number of failures per trial is, however, significantly smaller ( $p < .05$ ) for the heterogeneous condition. We have no adequate explanation for the slight superiority of the heterogeneous condition. One possibility which we have considered is that the heterogeneous condition reduced S's tendency to reproduce cue items instead of response items. The frequency of such cue intrusions was small and variable. For what it is worth, a count of the cue intrusions shows some greater frequency for the homogeneous condition, (55), than for the heterogeneous condition, (32).

Thus, the results not only fail to support Köhler's findings but tend to point in the opposite direction. The type of response bias which allowed us to account for the results of Experiments I and II could not operate in this situation, since S's responses were independent of the properties of color and size. The failure to find a difference in favor of the homogeneous pairs is, therefore, consistent with our interpretation of the

earlier experiments. It is not consistent with Köhler's perceptual hypothesis.

Our experimental results fail to provide any support for Köhler's hypothesis. It is possible to account for all the findings in terms of principles of associative learning. These facts do not in themselves constitute a disproof of Köhler's position. His hypothesis remains an hypothesis for which there is as yet no experimental evidence.<sup>6</sup>

*The weakness of perceptual analogies in the analysis of memory.* The difference between Köhler's analysis and ours again brings to the fore an issue which has long divided Gestalt theorists and associationists: the relation between perception and memory. The Gestalt theorist is concerned with the effects of perception on memory. Primary perceptual experience leaves a pattern of traces; the arousal of these traces mediates recall. In the analysis of learning and retention the Gestalt theorist is thus forced to maintain a precarious distinction between an initial experience (perception) and its later consequences (memory). It is difficult to specify where one ends and the other begins. Like the spurious present, the exact moment at which a primary perception changes into a memory or a perception-plus-memory must, we suspect, elude even the most sophisticated phenomenologist. Yet, the Gestaltist insists on looking for the influences of primary perceptions on memory. In his preoccupation with the persisting effects of phenomenal experience, he pays little attention, if any, to the

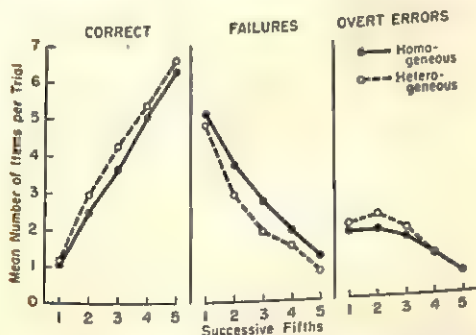


FIG. 4. Vincent curves for correct responses, failures to respond, and overt errors in Experiment III.

<sup>6</sup> We have not considered an experiment briefly mentioned by Köhler attempting to show that cues and responses that "fit together" are learned better than those that do not. We have not felt it possible to deal with this problem, since the obvious difficulties involved in an independent definition of "fittingness," as distinguished from familiarity, have not been resolved. Moreover, only a very sketchy account of the procedures and results is given.

conditions which influence *S's responses* from which inferences about both perception and memory must be made. Köhler's failure to consider the errors as well as the correct responses in relation to the characteristics of the series is an excellent case in point. Stimuli, "experiences," and "traces" cannot tell the whole story. There are always the responses that we must worry about.

The association theorist does not draw a sharp conceptual distinction between the initial "perception" and the recall tested at a later time. Rather, he is concerned with the *changing probabilities of different stimulus-response sequences* as a function of training. In paired-associate learning, for example, the correct response can be given initially only when the response item is exposed in the window of the memory drum ("perception"). Gradually, the cue items come to elicit the correct responses before the total pair has been exposed ("memory"). But the cue items also elicit incorrect responses ("memory?"). The changes in the probabilities of different responses vary systematically with the characteristics of the series and the conditions of practice.

The difference between this formulation and that of the Gestaltist is not purely semantic. The emphasis on con-

ditions of responding (*a*) avoids the methodological impasse of distinguishing between primary perceptions and memories, and (*b*) avoids a one-sided emphasis on sensory input at the expense of a careful analysis of the learner's behavior.

#### REFERENCES

1. GIBSON, ELEANOR J. A systematic application of the concepts of generalization and differentiation to verbal learning. *Psychol. Rev.*, 1940, 47, 196-229.
2. HEBB, D. O. *Organization of behavior*. New York: Wiley, 1949.
3. KÖHLER, W. On the nature of associations. *Proc. Amer. Philos. Soc.*, 1941, 84, 489-502.
4. MADSEN, A., & LENNOX, M. A. Response to colored light flash from different areas of optic cortex and from retina in anesthetized cat. *J. Neurophysiol.*, 1955, 18, 574-582.
5. MANDLER, G. Response factors in human learning. *Psychol. Rev.*, 1954, 61, 235-244.
6. MCGEOCH, J. A. The vertical dimensions of mind. *Psychol. Rev.*, 1936, 43, 107-129.
7. MCGEOCH, J. A., & IRION, A. L. *The psychology of human learning*. New York: Longmans, Green, 1952.
8. SKINNER, B. F. Are theories of learning necessary? *Psychol. Rev.*, 1950, 57, 193-216.
9. THORNDIKE, E. L. *Human learning*. New York: Century, 1931.

(Received April 25, 1956)



# THE PSYCHOLOGICAL REVIEW

## EFFECTS OF THE GESTALT REVOLUTION: THE CORNELL SYMPOSIUM ON PERCEPTION<sup>1</sup>

JULIAN E. HOCHBERG

*Cornell University*

*"Structuralism" and the Gestalt revolution.* Before the turn of the century, the psychology of sensation and perception enjoyed in "structuralism" a relatively unified approach. New data and interests soon forced the system apart at the seams; as to perception, the most important criticisms came from the Gestaltists. One main purpose of this symposium was to consider what areas of agreement could be found three decades after the introduction of *Gestalttheorie*. Discussion centered around two areas of research: perceptual change

(or learning), and the perception of "events." Let us review briefly what made these areas important to Gestalt theory.

*The "Constancy Hypothesis" and "Event Perception."* Structuralism tended to assume that: "... things look as they do because the proximal stimuli are what they are" (16, p. 80), and that "... the result of a local stimulation is constant ... that all locally stimulated excitations run their course without regard to other excitations" (16, pp. 96 f.)—the "Constancy Hypothesis." Any failure of this relationship was explained in terms of previous associations: "... things do not look as they ought to on the ground of pure perceived stimulation, and they differ from such an expectation by looking more like ... the things with which we have real dealings. ... in dealing with things we acquire experience about them, and this experience enters our whole perception" (16, pp. 84 f.).

It is easy to embarrass the Constancy Hypothesis. We can obtain the same responses to different stimuli or different responses to the same stimuli, as in the phenomena of object constancy, the geometrical illusions, etc. (16). Auditory stimuli affect visual experience, and vice versa (24); forms appear identical when stimulating to-

<sup>1</sup> Participants were: Egon Brunswik, University of California; James Drever, University of Edinburgh; James J. Gibson, Cornell University; Fritz Heider, University of Kansas; Julian Hochberg, Cornell University; Gunnar Johansson, University of Stockholm; George Klein, New York University; Ivo Kohler, University of Innsbruck; Robert B. MacLeod, Cornell University; Wolfgang Metzger, University of Munster; T. A. Ryan, Cornell University; Hans Wallach, Swarthmore College.

This symposium, held in June, 1954, was made possible by a National Science Foundation Grant. No attempt was made to preserve either the specific order of discussion, or the identity of the source of each opinion, though I have in a few instances inserted names of contributors. Since such reworking tends to take the responsibility out of the hands of the other participants, I have endeavored to return it to them by circulating copies of the report for their comments, addenda, and post-mortem alterations, and to take account of their replies, where length permitted.

tally diverse elements,<sup>2</sup> yet are rendered unrecognizable without changing the elements (16). The Gestaltists held that the theoretical error lay in the choice of too-small units of analysis; this was to be avoided by a "naive" phenomenology as opposed to formalized analytic introspection. Moreover, the "elements" are not independent, and the "laws" of their interaction were to be understood in terms of a new and unorthodox view of the underlying molar cerebral processes.

The phi phenomenon is a good example here. The (illusory) apparent motion can be made as completely convincing as real motion; the explanation of perceived real motion as the tying together by association of successive sensations of position becomes questionable. The apparent motion is a unitary experience, and Wertheimer started the Gestalt search for "brain models" by postulating an equally unitary cortical process. Perhaps more important, when Korte (18) investigated the dependence of the phi phenomenon upon brightness, separation, etc., a new form of psychophysics began to emerge, in which the experience was a full *event*, rather than a fragmentary "sensation," and in which *the stimulus variables manipulated were not necessarily ones which seemed intuitively similar to, or attributively responsible for, the quality of the response*. In general, Gestaltists hoped to restore stimulus-response correlations by treating entire configurations as the stimuli, and entire phenomenal events as the responses; in the process, the simple associational formulae of perceptual learning were discarded. Note that the Gestalt theorists

were *not* nativists in the traditional sense. They did not deny that learning affects perception. They denied only that all departures from the Constancy Hypothesis are to be explained in terms of past associations.

#### ADAPTATION TO THE DISTURBANCE OF STIMULATION: PERCEPTUAL RELEARNING

##### *Up-Down Inversion*

The effects of natural visual disturbances (long-term cataracts, etc.) have been reviewed (11), and contribute to a modern version of the old nativist-empiricist controversy; despite the recognized importance of the Stratton experiment (27), however, few experiments have been performed with humans (6, 27). In Erisman's laboratory, different kinds of protracted disturbance are under investigation, and Ivo Kohler's (cf. 17) findings are striking and important.

Let us survey the course of adaptation to visual disturbance, especially in inversion (in which up becomes down and vice versa).

*Stage 1.* When disturbing spectacles are first put on, the world seems strange in various ways: faces look unfamiliar, walking people seem mechanical (the up-and-down component, normally not "seen," becomes apparent), brightness contrast seems greater, colors more saturated. As head and eyes move, the normally stationary world swings about. The unfamiliarity of faces eventually wears off; in later stages, one can recognize features, etc., but even to the last, expressions cannot be discerned. The subject is almost incapacitated, with his motor actions appropriate at best to a world which appears upside down.

*Stage 2.* The subject can negotiate streets, can "fence" with the experimenter, etc. However, we cannot accept such task performance as the op-

<sup>2</sup> A recent experiment by Wallach and Austin (2) makes it necessary to qualify this statement, however, since it now appears that there is at least some dependence of form recognition on retinal locus of original presentation.

erational index of perception, or perceptual response (cf. 23), without residuum: *despite such effective motor performance, the world remains phenomenally inverted!* Perceptuo-motor adjustment, and phenomenal adaptation, run separate (but related) courses, and we cannot safely equate the two. Phenomenal invertedness first disappears for objects connected with the subject's body-system, or with clear indications of gravity. An object grasped, a plumb line, a face with a cigarette between its lips, with smoke ascending from the tip—these are seen as upright for a short time.

The relationship between the phenomenal upright, obtained through "adaptation" to an inverted stimulus, and the original phenomenal upright is by no means clear. Kohler reported that if a subject is shown two faces, one upside down (so that its retinal image is now, through the inverting spectacle, that which would without the spectacles have occurred from a face right side up) alongside an upright (and hence retinally inverted) face with a smoking cigarette between the lips, now *both* appear "upright"—*but in different ways and in opposite directions!* Kohler plans to adapt only *one* eye to inverting prisms, to permit further investigation of this question.

Stage 3. About a month later, the perceived world is almost continuously congruent with geography and independent head movements, even while the subject performs complicated behavior (bicycling, etc.). A specific familiar object (e.g., a given house) will appear unfamiliar when first viewed through the spectacles, even after other objects (e.g., other houses) have achieved familiarity through adaptation (Stage 3). Objects familiarized only through the spectacles become unfamiliar when the spectacles are removed. Transfer is restricted both for visual and motor

adaptation, and the process is not in any sense a learning to "invert the whole stimulus field."

Stage 4. If the spectacles are now removed, *the world appears inverted with normal stimulation.*

### Form Distortions

Adaptation to visual inversion may *not* be generalized, but adaptation to other disturbances may be. Wedge-deforming prisms contract one side of the field horizontally and expand it vertically, as compared with the other; phenomenal adaptation eventually occurs and, with spectacles removed, the reverse form distortion appears as an after-effect. Two classes of distortion effects appear: (a) general throughout the field, such as the *curvature* of lines; (b) conditional upon direction of gaze, such as the *angles* at which lines join.

Adaptation to width distortions (measured by rotation of an additional prism to eliminate phenomenal distortion) was slowest and weakest. Then came adaptation for curves, then for angles (at the right stage a subject will correctly see a square as *right-angled*, *but with curved sides*—an example of the inconsistency apparently permissible in the docile geometry of the visual field),<sup>3</sup> with adaptation to distortions of movement fastest and strongest.

Tentatively, therefore, it appears that adaptation may be faster when apparent differences between disturbed and familiar conditions are greater. Thus, adaptation is faster to up-down inversion than to right-left reversal, perhaps because our environment shows more bilateral symmetry around a vertical than a horizontal axis: gravity is a

<sup>3</sup> However, we cannot be sure from Kohler's report that this necessarily involves any visual inconsistency; if, as was not ascertained, the subject sees the square as being on an appropriately curved surface (Ryan), no inconsistency remains.



universal, whereas there are few comparable universals in a left-right sense.

### *Adaptation to Color-Fringes (Half Spectra)*

When viewed through prisms, each achromatic contour bears a fringe of color—a red-yellow half-spectrum to the left, blue to the right. This effect, also, is *general*, since the fringes appear everywhere in the field, independent of eye position. It is also *differential* in that, for any contour, the fringe is red-yellow on one side, and blue on the other.

After about a month, adaptation is complete: no more color-fringes are seen! *When the spectacles are removed, every contour in the field now appears surrounded by a reversed color fringe, i.e., blue on the side to which the red-yellow previously appeared, and vice versa.* If each eye adapts separately to opposite prismatic color fringes, each eye displays (in rivalry) its own after-effect. These data appear innocent enough, but are actually quite resistant to a satisfactory yet detailed explanation.

### *Split-Field Chromatic Disturbances*

The differential effects were more strikingly isolated as follows. With the left half of each spectacle lens blue and the right half yellow, white objects at first appear blue when viewed to the left, and yellow when viewed to the right. The usual desaturation appears if the eyes are fixed to one side but, obviously, as soon as they move to the other side, the complementary color appears. The astonishing fact is that, after protracted adaptation, eventual disappearance of all chroma occurs despite eye movement, in which the light at the fovea is at one time blue, at the next, yellow, and color is seen in neither case. With spectacles removed, *the aftereffect is also conditional upon eye*

*position: with the eyes left, the world appears yellow; with the eyes right, blue!* Thus, both effects are tied to, and therefore phenomenally independent of, eye movement.

### *Discussion of the Adaptation to Disturbed Stimulation*

Traditionally, for each sensation one could isolate one stimulus, and vice versa; perceptions entailed past associations, "unnoticed sensations," and "unconscious inference." The study of perception was possible only because, despite such associations, the underlying sensations remained tied to the stimulus and were rigorously investigable by the psychophysical methods. After Gestalt (and other) emphasis on the indistinguishability of sensation and perception, at least three extreme alternatives remained open:

1. To reconnect stimulus and percept by discovering the physiological mechanisms responsible for the (presumably) "organizational" lack of correspondence (cf. 16, 30); this has not, thus far, proved very fruitful.

2. To re-analyze the stimulus field in order to determine what aspects or "higher order variables," if any, *are* in correspondence with perceptual response (cf. Gibson, 9). This treats *perception* psychophysically as previously only *sensation* was investigated, and seeks to restore a new version of the Constancy Hypothesis.

3. To treat all stimulus-response relations as previously *perception* was investigated, i.e., as so determined by the organism's past experience with an unreliable environment that only the loosest, most statistical connection between proximal stimulus and perceptual response can be expected. Kohler's experiments constitute the greatest systematic change of relationship between stimulus and experience as yet reported,

and are well suited for examination of this issue.

Brunswik's approach comes closest to the third position: the correspondence between the response and the physical (distal) stimulus must be imperfect. This ambiguity, supposedly inherent in the equivocal relationship between distal stimuli and the proximal stimuli (the *cues*) to which the organism can respond, means that each cue has only a limited probability of being "correct." Thus, an *ecological sampling* reveals that the correlation is less than 0.6 between the cue of vertical position and the distal stimulus distance, and less than 0.4 between the cue of "space filling" (number of distinguishable intervening steps) and actual distance. An adjusted organism therefore must use many cues, weighting each according to its relative frequency (ecological validity) by some process of probability learning, of which Brunswik postulates two varieties: (a) *distribution learning*, and (b) *correlation learning*. Distribution learning (e.g., that hanging objects are rare) is very general, and, with enough reversals of cases, a "revolution" in such expectancies will occur. Correlation learning grasps concomitance, the reliability of any cue as a correlate of the object. Perception is, to Brunswik, a reasoning-like ("ratiomorphic") process, much faster to respond than is reasoning, but slower to change what has been learned; it is a speeded up, conservative, stereotyping, reasoning-like function (3).

The distributional perceptual learning provisions which are at the heart of this viewpoint may in principle accommodate Kohler's general findings, but they do not begin to explain the very singular specific data reported. Despite the apparent ease of the first Gestalt attack on the empiricist position, it is extremely difficult to refute any general

"explanation" couched in terms of past experience. Consequently, we must *specify* the relationship between past and present stimuli and percepts, or we have "explained" little more by invoking "past experience" than by invoking "human nature." A more specific attempt to account for the conditional effects through "conditioning" to the eye movements—or, more accurately, through extinction of previously conditioned responses (Drever),—did not seem very fruitful, although it may be possible to do more justice to findings with the greater number of assignable variables in the Hullian repertoire.

Metzger proposed as a Gestalt theoretical explanatory principle that *aspects of the visual world tend to become independent of behavior*, classing Kohler's results with the various perceptual constancies of size, motion, etc. (and explaining why reversed binocular disparity never adapted, since no added stability re the individual's behavior would thus be gained). However, this fails to explain adaptation to the half-spectra, in which the viewer's activity is hard to denote, and these phenomena seem at once the most general and the most puzzling. This proposal also leaves unexplained the fact that many aspects of the disturbance fail to reach complete adaptation, e.g., curvature adaptation appears asymptotic at about one-third. Metzger suggested that the "stresses" toward straightening are not strong enough to overcome completely the curvature of the actual stimulus; however, as Wallach pointed out, if the "internal stresses" of "organization" are insufficient to straighten an extreme curvature, adaptation should go farther with weaker prisms; yet Kohler reported that after prisms of a given strength caused an asymptotic degree of distortion-adaptation, prisms of double that strength cause adaptation to start



again, with a new asymptote, still short of complete correction.

Why not, Gibson asked, from a position he had previously taken on the problem of curved line adaptation (8), suppose that *the physical norm tends to become the phenomenal neutral point*? Such norms must still be defined,<sup>4</sup> but the consensus appeared to be, at this point, that a process of "partialing out" norms does seem to be a general principle during adaptation: what is *always* present becomes "unnecessary" to perceive. The blue half-spectrum to the right is always present, and so becomes the chromatic norm, and white light, the previous norm, will appear yellow. The problems affecting the previous formulations seem less troublesome from this point of view, but before its utility can be evaluated, much more explication is necessary.

Before we essay any conclusions, however, let us turn to the second general problem, that of *event perception*.

#### EVENT PERCEPTION

Recent advances have been made in three divergent salients—events involving the motion of several parts (Johansson, 15; Michotte, 22; Duncker, 5); the perception of social events (Heider and Simmel, 12); the kinetic depth effect, i.e., the depth perception consequent to sequential proximal stimula-

<sup>4</sup>Such norms need not be defined simply by statistical analysis of the distal stimulus. The straight line among curved lines, the rectangle among acute and obtuse angles, etc., are norms in other than a frequency sense: they are neutral points in "opposition series" (cf. 8), which run, say, from concave to the left to concave to the right, from obtuse to acute, etc. These are norms in a logical rather than an actuarial sense; a mathematical analysis of the visual world around us will be simpler, as Kohler puts it, if our reference tools are straight lines and right angles than if they are curves or obtuse or acute angles, so that the straight line and right angle would be analytic norms even if they never occurred in experience.

tion from a rotating distal stimulus (Wallach and O'Connell, 29; Metzger, 20).

#### *The General Problem*

We know that one can see motion when there is none on the retina. Johansson's series of studies (15) permit some fairly definite psychophysical relationships to be laid down for the perception of motion.

A horizontal row of four lights simultaneously moving up and down with the same phase appears as one entity in motion. With two pairs of lights 180° out of phase, "common fate" acts as a first, primitive law: those moving at the same time in the same direction appear as parts of one object. With more complex phase relations, those elements with the least phase difference form a group. Three-dimensional arrangements appear if they preserve a constant spatial distance between the elements in relative motion, picking out a constant shape in motion (cf. Metzger, 20, 21). When we depart from parallel harmonic motions to such motions at angles to each other ("Lissajous combinations"), the situation changes. With two moving lights having perpendicular paths, the same frequency, and phases such that they come to the point common to the two paths at the same time, they appear to move toward and away from each other *along the sloping line joining their extreme positions*. In addition, there appears a slight motion of the whole system in the perpendicular direction. If two points move in opposite directions with the same frequency in circular paths which, touching at one point, approach each other, they fuse, and then retreat on a horizontal line; simultaneously, the path along which they travel moves up and down. In short, *the physical motion which is present in the stimulus appears to have disappeared without any trace in per-*



ception, while a movement is perceived which appears to have no simple counterpart in the stimulus. However, the perceived motions *do* correspond to component vectors into which the stimulus motions may be analyzed: "... the phenomenon can fittingly be described as motions relative to two different co-ordinate systems, one fixed, and the other moving, in accordance with the principles for relative motion" (15, p. 97).

Thus, one can "extract" a given component from a complex motion and obtain the predicted "remainder." For example, a simple straight-line harmonic motion may be analyzed into two harmonic motions of lower amplitude and a 45° phase difference. Consequently, if one point is given an harmonic motion, and a second point is given the motion of one of its two components, the relative motion should be that of the remaining component; and in fact this is what is seen. One point moves back and forth, and the other moves back and forth with respect to it, like a planet rotating around its sun (in fact, although this *can* be seen as motion in a flat plane, many subjects see it in three dimensions, as an orbital motion).

In general, a "figural hierarchy" of motion is perceived: first, a static background; next, and in reference to this, the common motion; and highest, the components of motion relative to the common motion. If we have one point *a* moving in a circular path, and the other *b* in a vertical path, point *a* appears to move toward and away from *b* in a horizontal path, while both *a* and *b* move up and down. With *b* removed, *a* is seen to move in a circle; with *a* removed, *b* is seen to move vertically. If we repeat the procedure, but with the vertical path of *b* shorter than the diameter of *a*'s circular path, both points *a* and *b* move vertically together, with *a* describing an ellipse, and the vertical

component of *a*'s motion has lost the length of travel "used up" by *b*'s motion. In general, of two motions which are the same in other respects, the motion which has the shorter path or, more generally, the *lower velocity*, will determine the magnitude of motion of the combined system. It is this combined motion which forms the frame of reference within which the remaining motions occur. If two or more simultaneous motions or components of motion, equal in magnitude and direction, are analyzable within the stimulus motion, these will be seen as a single motion.

In general, Johansson states his findings as follows: In every case, if we abstract the motion components common to all of the moving points (as a special case, the common motion may reduce to zero), the remaining components become the relative motion of the parts, while the common motion becomes the motion of the whole relative to the stationary background. It is therefore always the shortest excursion which will determine the common motion.

#### *Motion and Three-Dimensional Space: The Kinetic Depth Effect*

In many of Johansson's experiments, the percept was three-dimensional, even though both proximal and distal stimuli were two-dimensional.

Just as we can study the psychophysical relationships in apparent motion without *necessary* concern with ontogenetic or historical antecedents, so can the emergence of depth as an aspect of event perception be studied in its own right. In fact, specific knowledge concerning perceptual learning begins to emerge from such study, to replace gross appeals to empiricism.

Particularly informative here are the investigations of the "kinetic depth effect" (29). Wallach was originally concerned with depth perception without

the "primary" depth cues (binocular disparity, etc.). In an at least partially empiristic view, one originally had "depth" experience only where there were such "cues" (making the case of monocular individuals awkward to explain); in terms of the "spontaneous organization" of *Gestalt-theorie*, we are just lucky, phylogenetically speaking. However, as one walks around in the world of objects, the proximal stimuli projected by distal objects are pretty much those which would be cast by the axial rotation of those objects, and the rotation of most of the objects around us will yield three-dimensional experience even when viewed monocularly. Can subjects (monocular or otherwise) "learn" from this kinetic effect, while in motion, the three-dimensional nature of a given object? (Remember also the importance of motion in the "relearning" of space in Kohler's findings.)

Wire patterns were chosen whose projections evoked two-dimensional responses; these gave way to three-dimensional responses when the wire forms were rotated. After this experience with a given form, it would tend in the future to appear three-dimensional, even when stationary. (This is not simply *knowledge* that the given form is projected by a three-dimensional object, since it displays unexpected and involuntary perspective reversal: what was near now looks far, with appropriate consequent size and shape "distortions.") Wallach terms this a *memory*, rather than *learning*, since only a specific "trace" is involved, and this is a long way from the creation of "trace complexes," with high generality of function—e.g., size constancy, which cannot be referred to any single memory trace.

But *why* should the rotation of a form bring its projected shadow perceptually into three dimensions?

The stationary two-dimensional pro-

jection of *any* given object is ambiguous in that there is an exceedingly large number of three-dimensional objects which will produce that projection. The *transformations* undergone by the projections as the objects rotate are unique, at least for many objects—but only if certain restrictive assumptions are made, e.g., that the projecting object must be a rigid form, etc. It is easy to say to this that, with the exception of animate creatures, most objects in our environment *are* rigid, and that this is simply a matter of distribution learning, in Brunswik's sense. However, we do not yet have any estimate of the ecological distribution of rigidity; moreover, Johansson finds with complex motions that rigidity is *not* always what is seen, even if the distal stimulus is really rigid (15).

Ambiguity (between distal and proximal stimuli) therefore must remain, but each successive momentary projection of the rotating object must, as Wallach points out, bring new information about it, and decrease the number of forms it may have and still give rise to the specific projection series. This does not immediately solve the problem of a working psychophysical correspondence, since any finite decrease of an infinite number of alternatives cannot do much toward specificity. However, if we note with Gibson that the transformation *sequence* undergone by the projection becomes *one* piece of information in the case of the usually "correct" perception of the distal stimulus, and that this requires less information to be specified than would that of describing each stage in the sequence separately—and if, as this writer has suggested elsewhere (13), that perceptual response will tend to occur which is the most economical of information—then a form of specificity returns to the stimulus-percept relationship. We must deal not with unique relationships between distal stimulus, proximal stimulus, and re-



sponse, but with relationships of varying probabilities. In Brunswik's terms, ecological validity and reliability may frequently be quite low. However, for most *normal* situations only one ecologically probable distal stimulus can yield the given proximal stimulus—if we choose adequate higher-order variables of stimulation for our analysis. Between proximal stimulus and perceptual response, then, there may be or may not be ambiguity of relationship. Where ambiguity is low, there is no problem for psychophysics; where ambiguity is high, a more probabilistic psychophysical correspondence may be in order.

A psychophysics without a punctate Constancy Hypothesis, and which employs relatively large units of analysis and tolerates probability statements of psychophysical relationship, permits the study of events of much wider scope than those we have been considering. The study of what may be termed the psychophysics of perceived causation, by Michotte (22), is only one step up the ladder from the event perceptions of motion and depth toward the study of the perception of self and others (7, 19, 25, 23). There is no intrinsic reason why such a "global psychophysics," to use MacLeod's term, cannot be applied to the study of any phenomena in which discriminational responses are obtainable. The utility of such research, of course, awaits empirical findings.

### *The Perception of Social Events*

A start has been made in Heider's pioneer researches on the "social perception" of moving geometrical figures (cf. 12). We have seen that perceived motion might be more closely determined by change of brightness than by actual motion; here, the Constancy Hypothesis suffers some further interdimensional fracturing: objective motion

may serve as stimulus not for perceived motion, but for perceived purpose, life or animation. Under Kohler's disturbed visual conditions, animate actions became mechanical motions; contrariwise, motion configurations may give rise to the perception of action and purpose, rather than of mechanical motion (15).

Heider's "animated cartoon" had a large triangle, a small triangle, and a small circle engage in some charmingly social peregrinations through the familiar "apparent motion" of stroboscopic presentation. Naive observers describing what they had just seen differed as to content, but agreed as to the actions in many portions of the film. Is this film only a projective technique, only a focal point around which individual antecedent experiences determine the response? If so, can at least gross similarities of experience be assumed or established? Answers can only be obtained by psychophysical research. However, we can appeal to certain at least peripherally relevant considerations.

Klein reported that when Heider's film is shown to psychotic patients (paranoid), the contents differ from those reported by normals, but the actions seen are the same. Despite the probably very great differences from individual to individual, certain stimulus variables seem important; many of us who watched Heider's film felt that, in general, uniform motion was not a compelling stimulus for perceived action, depending more upon context for its animate qualities than did nonuniform motion. Johansson reported that rows of lights moving with simple phase relations (say  $180^\circ$ ) were described by subjects in terms of some moving mechanism; with more complex phase relations (say  $90^\circ$ ), the motion was that of a wave, snake, etc.; with still more complex phase relations (say  $45^\circ$ ,  $135^\circ$ , and  $45^\circ$  between the lights), almost always living motion. Tentatively, then,



perhaps nonrigid, nonuniform motion is conducive to the perception of action.

### SUMMARY AND CONCLUSIONS

#### *Empiricism-Nativism*

Empiricist and nativist "explanations" were essayed quite frequently throughout the discussion, and proved equally unedifying. What one participant considered stimulus-determined (and probably innate) another considered at least partially "learned" (e.g., compare 14 and 28). One suspects that the issues of this traditional controversy have been reduced to specific questions of conditions and mechanisms which have to be settled by detailed investigation. Perhaps this is due to the fact that (at least with humans) it appears possible to find an empiricist "explanation" for almost any perceptual phenomenon; however, without concrete knowledge about perceptual learning, such general accounts are quite useless, if well-nigh invulnerable.

#### *The Minimum Principle*

The most generally acceptable summing up of Kohler's results was as follows: Two kinds of association between distal and/or proximal stimuli (including tactual-kinesthetic feedback stimulation) in our visual environment, can be at least roughly distinguished, namely invariable relations ( $r \cong 1.0$ ) and contingent relations ( $r < 1.0$ ). Where an invariable relationship occurs, instead of repeating it in each perceptual response, it becomes "partialled out" as norm, framework, or neutral point. As Kohler finally put it, it is as though the mathematical description of the world made by the organism takes the simplest possible form.

In Johansson's motion studies, those components in a complex moving stimulus which are common to all members of a group are "partialled out" and form a single framework in relation to which

the residual motions appear. Such unification achieves an "informational" economy since, for any given stimulus, the percept entailing the least number of changes is obtained. Objectively changing color or brightness tends to be perceived as appropriate changes in the motion of objects of unchanging color or brightness.<sup>5</sup> In general, wherever a response in terms of a single unchanging distal stimulus moving in depth would be more "economical" than one in terms of continually changing relationships in two dimensions, subjects will tend to report the three-dimensional alternative. Wallach has shown that this kinetic stimulus for depth can endow even motionless projections of the distal stimulus with perceived depth, so that it seems likely that we have here one of the building-blocks by which perceptual space is achieved or modified by perceptual learning.

With respect to the perception of social events, little is yet known aside from the demonstration by Heider (and incidental observations by Michotte) that psychophysical study of such higher-order variables of stimulation seems both possible and fruitful. Heider feels that we cannot talk about those

<sup>5</sup> Johansson finds that if the brightness of two motionless spots is varied with the same frequency, but 180° out of phase, the perceptual response quickly stabilizes as "a light of constant brightness moving back and forth behind the two windows." With a colored filter behind one window, a moving *white* light of constant intensity is perceived to move with no alteration, and the *window* is seen as colored. If color changes synchronize with the intensity changes, a light seems to travel back and forth behind the two windows, which are changing colors as though a bicolored curtain were being raised and lowered. That is, it seems preferable to see two such separate motion events rather than a single one which involves both motion and color changes. This appears to be quite general. Changes of stimulation which are objectively changes along other dimensions—brightness, color, etc.—tend to be ascribed to changes in the dimension of spatial motion.

motions which are perceived as actions except in terms of short, readily remembered concepts, and that the simplest action will determine the perceptual response (cf. Brunswik's "ratio-morphism"); thus, those motions most difficult to describe in simple mechanical terms appear most readily as actions.

### *Residual Effects of the Gestalt Revolution*

After the elimination of those principles upon which agreement could not be reached, several important points remain.

1. It is frequently useful to admit subjects' statements as to what they see ("naive" phenomenal reports) as at least contributory evidence about the percept (e.g., consider the separate courses taken by perceptuo-motor adaptation and phenomenal adaptation to visual disturbance). We cannot identify percept with perceptuo-motor manifestations, despite recent trends to do so. This is not an epistemological issue, but simply one of methodology.

2. Attention should be focused on higher-order variables of stimulation. The attempt to return to the restricted (and once presumably physiologically identifiable) "atomistic" units of pre-Gestalt days seems hopeless. Whatever measurable aspects of stimulation over space and time may be extracted and brought into correspondence with abstractable dimensions of response now constitute fair game for the investigator of perception.

3. Of all of the positive Gestalt formulations which sought to replace the Constancy Hypothesis, the principle of "simplicity" or "maximum homogeneity" has proved most general. It seems as quantifiable in terms of the objective stimulus as most of the others, and more so than some, such as the "law" of *Prägnanz*. Since existing physiological models fail to explicate the "laws" of sensory cortical organization to any use-

ful extent, the necessity for predicting response from stimulus characteristics has become progressively more urgent. The precise form of the statement and the unequivocality of the minimum principle varied with the discussants but, in one form or another, it was held by all present. What to me is its most promising approximation to date—the formulation that, other things equal, that perceptual response to a stimulus will be obtained which requires the least amount of information to specify (1, 13)—obtained agreement among Heider, Gibson, Kohler, Metzger, Johansson, and myself. Some differences existed as to whether or not this tendency is itself likely to be learned, or how precisely it should be formalized; at its simplest, it could be described as a "laziness of the perceptual imagination" (cf. Wheeler's "law of least action," etc.).

Several areas stand out as important for future investigation.

1. The problem of perceptual learning remains very poorly understood or formulated, and yet is an extremely critical area (cf. 10).

2. The perception of space, depth, and distance is frequently treated in the textbooks as a solved problem. Despite the fact that some restricted areas of precise and applicable knowledge exist, however, the basic problems in this area are completely *unsolved*, and we must launch a fresh attack on what is historically one of the oldest of the systematic problems of psychology.

3. The perception of physical and social events is an area of great promise not only for the field of perception but for potential application within and without psychology. However, it must be confessed that—aside from a very few pioneer studies in event perception, facial expression of emotions, the so-called "physiognomic perceptions," and some specific esthetic and artistic investigations (e.g. 4, 12, 25)—we know

little more about the general area than that fruitful research seems possible.

4. The study of the results of prolonged visual disturbance is critically important in the understanding of the perceptual process.

5. The *minimum principle* (the "principle," in the traditional sense, which met with most general agreement) requires more concrete and self-conscious research tests, to determine precise applicability and limits, and to avoid what Wallach and Johansson see as a potential unwarranted elevation to a meta-physical dictum.

# REFERENCES

1. ATTNEAVE, F. Some informational aspects of visual perception. *Psychol. Rev.*, 1954, 61, 183-198.
2. AUSTIN, P., & WALLACH, H. The effect of past experience on the reversal of ambiguous figures. *Amer. Psychologist*, 1953, 8, 314. (Abstract)
3. BRUNSWIK, E. "Ratiomorphic" models of perception and thinking. In *Proc. 14th Internat. Congress Psychol.*, Amsterdam, North Holland, 1955. Pp. 108-110.
4. BRUNSWIK, E., & REITER, L. Endruckscharaktere schematisierter Gesichter. *Z. Psychol.*, 1937, 142, 67-134.
5. DUNCKER, K. Über induzierte Bewegung. *Psychol. Forsch.*, 1929, 12, 180-259.
6. EWERT, P. H. A study of the effect of inverted retinal stimulation upon spatially coordinated behavior. *Genet. Psychol. Monogr.*, 1930, 7, 177-363.
7. GAGE, N. L. Accuracy of social perception and effectiveness in interpersonal relationships. *J. Pers.*, 1953, 22, 128-141.
8. GIBSON, J. J. Adaptation with negative after-effect. *Psychol. Rev.*, 1937, 44, 222-244.
9. GIBSON, J. J. *The perception of the visual world*. Boston: Houghton Mifflin, 1950.
10. GIBSON, J. J., & GIBSON, E. J. Perceptual learning-differentiation or enrichment? *Psychol. Rev.*, 1955, 62, 32-41.
11. HEBB, D. O. *The organization of behavior*. New York: Wiley, 1949.
12. HEIDER, F., & SIMMEL, M. An experimental study of apparent behavior. *Amer. J. Psychol.*, 1944, 57, 243-259.
13. HOCHBERG, J., & McALISTER, E. A quantitative approach to figural "goodness." *J. exp. Psychol.*, 1953, 46, 361-364.
14. HOCHBERG, J. E., & BECK, J. Apparent spatial arrangement and perceived brightness. *J. exp. Psychol.*, 1954, 47, 263-266.
15. JOHANSSON, G. *Configurations in event perception*. Uppsala: Almqvist & Wiksell, 1950.
16. KOFFA, K. *Principles of Gestalt psychology*. New York: Harcourt Brace, 1935.
17. KOHLER, I. Über Aufbau und Wandlungen der Wahrnehmungswelt. *Oesterr. Akad. Wiss. Philos.-Histor. Kl.; Sitz.-Ber.*, 1951, 227, 1-118.
18. KORTE, A. Kinematoskopische Untersuchungen. *Z. Psychol.*, 1915, 79, 193-296.
19. MACLEOD, R. B. The place of phenomenological analysis in social psychological theory. In J. H. Rohrer & M. Sherif (Eds.), *Social psychology at the crossroads*. New York: Harper, 1951.
20. METZGER, W. Tiefenerscheinungen in optischen Bewegungsfeldern. *Psychol. Forsch.*, 1935, 20, 195-260.
21. METZGER, W. *Gesetze des Sehens*. Frankfurt: Kramer, 1953.
22. MICHOTTE, A. *La perception de la causalité*. Louvain: Inst. Sup. de Philosophie, 1946.
23. POSTMAN, L. Experimental analysis of motivational factors in perception. In *Current theory and research in motivation*. Lincoln: Univ. of Nebraska Press, 1952.
24. RYAN, T. A. Interrelations of the sensory systems in perception. *Psychol. Bull.*, 1940, 37, 659-698.
25. SCHLOSBERG, H. The description of facial expressions in terms of two dimensions. *J. exp. Psychol.*, 1952, 44, 229-237.
26. SNYDER, F. W., & PRONKO, N. H. *Vision with spatial inversion*. Wichita: Univ. of Wichita Press, 1952.
27. STRATTON, G. M. Vision without inversion of the retinal image. *Psychol. Rev.*, 1897, 4, 341-360; 463-481.
28. WALLACH, H. Brightness constancy and the nature of achromatic colors. *J. exp. Psychol.*, 1948, 38, 310-324.
29. WALLACH, H., & O'CONNELL, D. N. The kinetic depth effect. *J. exp. Psychol.*, 1953, 45, 205-217.
30. WERTHEIMER, M. Untersuchungen zur Lehre von der Gestalt II. *Psychol. Forsch.*, 1923, 4, 301-350.

(Received May 1, 1956)



## THE ROLE OF PRIMARY DRIVE REDUCTION IN FIXATIONS<sup>1</sup>

ROBERT S. FELDMAN

*University of Massachusetts*

Maier and his students have demonstrated that subjecting rats to insoluble situations on a Lashley jumping stand leads to behavior rigidity, and that this rigidity, characterized by fixations, persists even when more adaptive responses are available to the animal (4). Moreover, the fact that animals that have been in insoluble problem situations form a bimodal distribution of learning measures during a subsequent soluble problem led Maier (5) to propose that when conflict occurs it leads to behavior that is compulsive in character, in that it is not influenced by consequences such as selective reward and punishment. Others, however, feel that Maier's evidence can or should be incorporated within a rubric of motivational and learning laws and hypotheses. For example, Mowrer (10) has suggested that the reinforcing effect of anxiety reduction is the key idea, while Miller (9) holds that the motivational force of an acquired drive of fear accounts for the compulsive persistence characteristic of behavior associated with trauma of one sort or another. These points of view were subsequently discussed by Maier and Ellen (6) who found them inadequate to account for all of their evidence relating to fixations.

This paper deals primarily with another attempt made by Wolpe (12) to account for Maier's evidence in terms of learning principles. Referring to Maier and Ellen's criticism of the anxiety-reduction explanation, Wolpe states:

There is little doubt that the experimental facts they [Maier and Ellen] give do invali-

date the anxiety-reduction hypothesis. But . . . the same facts are quite in keeping with the differently formulated hypothesis. . . . Each time a jump is forced by the air-blast, it is reinforced by reduction of the air-blast-induced drive. This is a primary drive, and its reduction is clearly overwhelmingly responsible for the reinforcement; for the platform situation is at no stage conditioned to a sufficient level of secondary drive to be able to impel jumping in the absence of the air-blast. When jumping in a particular way has thus been repeatedly reinforced, it becomes firmly established as the habitual response to the air-blast stimulus, and the more firmly it is established, the weaker does the competing alternative response tendency become (12, p. 114).

Familiarity with Maier's data suggests that this hypothesis can be questioned on two counts. First is Wolpe's statement that "the platform situation is at no stage conditioned to a sufficient level of secondary drive to be able to impel jumping in the absence of the air-blast." This statement seems to say that every response is caused by air blast, but this is not true at all. As a matter of fact, after a few days of conflict trials, very few trials require the application of air blasts or shocks to force a response in an insoluble or soluble problem situation. But this is a minor point. The more important point is that Wolpe held that the reason the rats do not give up their fixated responses for adaptive responses is that the stimulus to jumping is the air blast and not the stimulus cards. If it is true that air blast is an important variable in fixations, a simple way to find support for this possibility would be to compute both the number of times a rat was given air blast and the total duration of air blast during frustration trials, and compare these scores with some indicator of response strength.

<sup>1</sup> This research was supported by a research grant (MH-376) from the National Institute of Mental Health, National Institutes of Health, United States Public Health Service.

In an experiment by Maier and Feldman (8) three split-litter groups of rats were placed in a frustrating, insoluble-discrimination problem for different numbers of trials: one group for 80 trials, a second for 160 trials, and the third for 240 trials. All rats received 10 trials per day, and were driven by air blast if they failed to make a response within 30 seconds. This procedure generally led to persistent position responses, but if a rat developed a preference for one of the cards it was discarded, since only position-responding animals were desired for the experiment. As controls, three other groups were *trained* to form position responses by locking the window in one position 100 per cent of the time and unlocking the opposite window; one group practiced the response for 80 trials, the second for 160, and the third for 240 trials. The strength of the position response for all rats was then measured for its resistance to extinction when the problem was changed to require the learning of a card discrimination in the same apparatus. During this stage the experimenter forced the rats to respond to the correct window on trials alternated with free trials by placing his hand alongside the rat and nudging it toward the correct window. This technique insured that all rats would eventually abandon the position responses and learn the discrimination.

The measure indicating that the position response had been abandoned was the number of that trial during which the rat first makes a response different from its position response during a free trial. This number is designated as a breaking score. The results showed that the three control groups which had been *trained* to form the position responses gave up the position responses and adopted the card responses at the same rate. Thus, the duration of the practice period was not a significant variable.

The three experimental groups, however, gave up their frustration-instigated position responses after many more trials than it took for the controls; the critical ratio of the difference was in excess of six. On the other hand, the mean learning scores, i.e., scores indicating that the discrimination had been mastered, did not vary significantly for any of the groups, experimental or control.

It was seen from these results that the conflict groups were different from the trained groups in that the number of trials necessary for abandoning the position response was greater when it was acquired during frustration than when it was learned under conditions of motivation. Also, since the mean learning scores were approximately the same for all conflict and all motivated groups, it is easily deduced that the conflict groups needed significantly fewer additional trials after breaking to reach the learning criterion than did the motivated groups. This result suggested that learning took place in the discrimination problem for the conflict animals while they still expressed the position response, but that the old position stereotype<sup>2</sup> was too strong to permit the newly acquired response to come into expression.

Since the foregoing analysis assumed that the strength of a stereotype could be measured by the number of trials needed to break it, Wolpe's hypothesis would also be supported if there was a positive relationship between air-blast scores and breaking and learning scores. Therefore, product-moment correlations were computed between air-blast scores (the number of trials during the training and conflict stages during which Ss

<sup>2</sup> Attention is called to the distinction between a stereotype and a fixation. A stereotype is a consistent response to a position or to one of the windows during an insoluble problem situation. A fixation is a consistent response to a position or window when a more adaptive response is available.



## PRIMARY DRIVE REDUCTION

TABLE 1  
TOTAL NUMBER OF AIR BLASTS CORRELATED WITH LEARNING AND BREAKING SCORES, AND TOTAL DURATION OF AIR BLAST CORRELATED WITH LEARNING AND BREAKING SCORES

Group	N	Trials with Air Blast vs. Learning Scores	Trials with Air Blast vs. Breaking Scores	Duration of Air Blast vs. Learning Scores	Duration of Air Blast vs. Breaking Scores
8-day	37	-.23 (.14)*	.32 (.11)	-.07 (.15)	-.12 (.18)
16-day	37	.005 (.17)	.16 (.14)	.28 (.02)	-.07 (.17)
24-day	37	.13 (.16)	.12 (.16)	.15 (.16)	.19 (.16)

\* The standard errors of the correlations are in parentheses.

received air blast) and learning scores, and between air-blast scores and breaking scores. Also, correlations were computed using total air-blast duration as the air-blast variable. Since it is desired to test the hypothesis that air blast is related to conflict-induced stereotypes, this analysis was confined to the experimental groups. The correlations with their standard errors are shown in Table 1. It is seen that practically all of the coefficients are of low order, the highest one being .32 ( $\sigma_r = .11$ ); the reliabilities are also low. This would indicate that there can be virtually no prediction as to the persistence of a conflict-induced stereotype from the number of times an animal was subjected to air blast, or from the total duration of air blast. Since these results may reflect the possibility that rigid avoidance reactions to air blast were established after few trials and that the persistence of these reactions reached maximum levels after few more, the data of the 8-day group might be scrutinized more closely. The animals in this group received the least amount of air blast, thus minimizing the possibility that rapid and strong avoidance conditioning had occurred. Therefore, if any trend in support of Wolpe's hypothesis could be found, the animals in the 8-day group would be most likely to reveal it. But the 8-day group's relatively unreliable coefficient of .32, for the relationship between the number of trials with air blast and breaking scores, is the only

one in Table 1 that could even remotely support Wolpe's contention, whereas all the other coefficients show a trend in the direction opposite to that demanded by the hypothesis.

Furthermore, it may be recalled that in many frustration studies by Maier and others, grid shock rather than air blast was used to force responses. To explore fully the possibility that fixated responses are conditioned avoidance responses, comparisons were made between groups of rats that fixated responses and those that did not, in terms of their shock scores, viz., the number of trials accompanied by shock during the period when the responses were initially established.

In a study by Neet and Feldman (11),<sup>3</sup> 72 rats were trained to respond on a Lashley jumping stand and then subjected to an insoluble problem for 160 trials. During a 40-trial training period and during conflict trials, Ss that failed to jump within 30 seconds were forced by applying shock to a floor grid on the jumping platform. Following the insoluble problem, all Ss were pre-sented with a soluble brightness-discrimination problem. Testing continued until Ss made no more than one error

<sup>3</sup> The data for this analysis were taken from a study by Neet and Feldman (11) that investigated the effect of electroconvulsive shocks on fixated responses. Their results showed that a series of 10 or 25 daily ECSs had no appreciable effect on the stability of the fixated response.



in three successive days of testing, or until they had a total of 200 test trials. The results showed that only 12 rats were able to abandon the conflict-induced stereotypes and meet the learning criterion within a limit of 200 trials. The mean number of trials that were accompanied by shock during the training and insoluble problem stages was computed for the Ss that solved the subsequent soluble problem and for those that did not within the 200-trial limit. It was found that the solvers received grid shock during a mean of 66.2 trials and the nonsolvers during a mean of 41.2 trials. This difference was significant ( $p = .05$ ), but the difference was in the opposite direction to that demanded by Wolpe's hypothesis, that is to say, the Ss that received more shock had weaker stereotypes, judging at least by their ability to alter them.

The foregoing analysis points to the difficulty of relating the rigidity and the persistence of fixations to what appear to be obvious primary or secondary reinforcers within the insoluble problem situation described by Maier. It will be recalled that Wolpe argued quite specifically that the frequency of application of the driving stimulus (in his case, air blast) is related to the strength of the response. He stated, "Increased air-blast means increased drive with consequent greater drive reduction" (12, p. 114). Yet in our analysis of air blast and grid shock this relation does not hold at all, if one interprets "greater drive reduction" as being the primary condition for response rigidity.

Other studies point to the difficulty of specifying cues and reinforcers that relate to fixations. Feldman (1, 3) has shown that the tendency to fixate responses is not tied to discriminative cues that seem to direct them, e.g., on a Lashley jumping stand, when rats had fixated *jumping* responses to one of the cards and when on alternate trials they

were permitted to walk along runways to the windows, some of them would develop *walking* fixations to a position. Other rats fixated position responses on *jumping* trials and card responses on *walking* trials, and a group of rats that were given alternate jumping and walking trials during a solvable discrimination problem learned to walk to the correct window but persisted in a jumping fixation for a total of 200 trials. Feldman (2) has also shown that even when practice for a stereotyped response is controlled, by forcing rats to make every possible type of response in an insoluble problem situation an equal number of times, fixations still develop in subsequent discrimination problems. Maier and Ellen (7) recently demonstrated that more stereotyped responses induced by 50 per cent random reward and punishment were abandoned, and fewer fixations occurred, during a subsequent soluble problem when the correct card appeared on the fixated side 50 per cent of the time than when it appeared on that side 80 per cent of the time, even though the latter reward schedule should be more reinforcing and contribute to more learning.

Finally, consideration is directed to the common-sense argument that the rat in the insoluble problem situation has simply learned that a response to a position is as good as any other response, even though it is rewarded only 50 per cent of the time; and that if the response is made with anticipation of failure by hitting the window with the shoulder instead of the nose, punishment can be minimized. The argument further supposes that during the subsequent soluble-discrimination problem the position stereotype persists, and that the rat never tries the opposite side because that side is more likely to be more punishing since the rat has not learned to avoid bumping its nose on that side.

This explanation essentially suggests

primary reinforcing mechanisms to account for behavior rigidity, but there are two shortcomings in this approach. First, one cannot account for fixations *to one of the cards* on this basis, since rats jump left and right with equal facility, but always to the incorrect card. It is emphasized that in our experiments, because of barriers adjacent to the windows, rats with *card* fixations cannot jump so abortively as to miss the card altogether; thus they "try" and fail on every trial. It would make no sense, therefore, to say that these rats have a response that is less punishing because it is the *most* punishing, yet rats with a history of conflict trials usually do not attempt any other response.

Second, the foregoing explanation fails to account for the manifest difference between rats that acquire a conflict-induced position response and those that learn a position response by being rewarded on one side and punished on the other in a consistent fashion. In the latter case, rats learn to make the best response during position learning, but readily abandon it when changed conditions call for a response to one of the cards; the first solution has not precluded the second. If it is assumed that during the insoluble problem the position stereotype is the best adjustment possible, it is not clear why in this case this solution renders the animal incapable of solving subsequent problems. This is the crux of the matter.

#### SUMMARY

The data of two experiments are re-analyzed to find support for Wolpe's contention that fixations are the consequence of primary drive reduction. In one experiment, using the Lashley jumping-stand technique, three groups of rats were submitted to an insoluble problem situation for 8, 16, and 24 days, respectively. The Ss were then given a soluble discrimination problem in which they

were forced to respond to the correct window on every other trial. During all procedures, the rats were motivated to respond by applying air blast if they hesitated for more than 30 seconds during any trial. Product-moment correlations were computed for the relationship between air-blast variables (number of conflict trials accompanied by air blast, and total durations of air blast), and criteria for abandoning the conflict-induced stereotypes and learning the discrimination problem. There was no correlation of sufficient magnitude to permit a prediction of the strength of the conflict-induced response from the amount of air blast. In a second experiment, a group of 72 rats was submitted to an insoluble problem as in the first experiment. Only 12 rats were able to solve a subsequent soluble problem. During all procedures, rats were motivated to respond by grid shock if they hesitated for more than 30 seconds on the jumping platform. Comparing the frequency of shock between the Ss that abandoned the conflict-induced stereotype and solved the problem and those that failed, it was found that the rats that *solved* had significantly more shock during the conflict trials. These results were considered as failing to support the notion that fixations are the consequence of the reduction of drive induced by air blast (or shock).

#### REFERENCES

1. FELDMAN, R. S. The relationship between guidance and the specificity of the fixated response in the rat. Unpublished doctor's dissertation, Univer. Microfilms, Publ. No. 1403, Univer. of Michigan, 1950.
2. FELDMAN, R. S. The role of response habituation in response fixation. *Amer. Psychologist*, 1953, 8, 348. (Abstract)
3. FELDMAN, R. S. The specificity of the fixated response in the rat. *J. comp. physiol. Psychol.*, 1953, 46, 437-492.
4. MAIER, N. R. F. *Frustration: the study of behavior without a goal*. New York: McGraw-Hill, 1949.

5. MAIER, N. R. F. The premature crystallization of learning theory. In *Kentucky Symposium: learning theory, personality theory, and clinical research*. New York: Wiley, 1954. Pp. 54-65.
6. MAIER, N. R. F., & ELLEN, P. Can the anxiety-reduction theory explain abnormal fixations? *Psychol. Rev.*, 1951, 58, 435-445.
7. MAIER, N. R. F., & ELLEN, P. The effect of three reinforcement patterns on positional stereotypes. *Amer. J. Psychol.*, 1955, 68, 83-95.
8. MAIER, N. R. F., & FELDMAN, R. S. Studies of abnormal behavior in the rat: XXII. Strength of fixation and duration of punishment. *J. comp. physiol. Psychol.*, 1948, 41, 348-363.
9. MILLER, N. E. Studies of fear as an acquirable drive: I. Fear as motivation and fear-reduction as reinforcement in the learning of new responses. *J. exp. Psychol.*, 1948, 38, 89-101.
10. MOWRER, O. H. *Learning theory and personality dynamics*. New York: Ronald, 1950.
11. NEET, C. C., & FELDMAN, R. S. The effect of electroconvulsive shock on fixated behavior in the rat: I. The effect of a ten- and of a twenty-five-day series of ECS on the stability of the fixated response. *J. comp. physiol. Psychol.*, 1954, 47, 124-129.
12. WOLPE, J. Learning theory and "abnormal fixations." *Psychol. Rev.*, 1953, 60, 111-116.

(Received February 6, 1956)



## ANALYSIS OF EXPLORATORY, MANIPULATORY, AND CURIOSITY BEHAVIORS

WILLIAM N. DEMBER<sup>1</sup> AND ROBERT W. EARL

University of Michigan

A class of behaviors apparently not motivated by the commonly accepted homeostatic drives, nor by learned accretions to these drives, has come under considerable experimental investigation in recent years. Such behavior has been variously labeled exploration, manipulation, or curiosity. Various labels have been applied to the conditions for arousal of this behavior, chief among them novelty and complexity. While this behavior has been discussed at length by Nissen (16) and Harlow (8), there is available as yet no comprehensive theory to account for it. Hebb's "short-circuiting" (9) has been offered as an explanatory concept at the neurophysiological level. It is our purpose in this paper to suggest a psychological structure into which such behavior and the conditions of its arousal might fit.

### CHARACTERISTICS OF THE BEHAVIOR

In any attempt at classification of events there are two dangers: (a) the inclusion in the same category of events which only superficially belong together, and (b) categorization as different events those which are only superficially dissimilar. In this paper we are more likely to be guilty of the first error.

We would assert that those behaviors labeled "exploration," "curiosity," etc., belong to the general class of behavior, *attention*. Functionally, this behavior brings the organism into contact with certain portions of its environment rather than others. Such behavior differs from attention of the textbook variety only in the degree of locomotor involvement. One may attend to a loud noise without very gross motor ad-

justments, whereas a rat, to make contact with the end box of a T maze, ordinarily must get there on its own four feet.

This point of view enables us to consider in a single system such diverse phenomena as binocular rivalry, manipulatory behavior of the monkey, exploratory activity of a rat in a maze, and certain aspects of the aesthetic and intellectual behavior of human beings. An obvious advantage of this scheme is that it affords a multitude of situations for inducing and testing hypotheses, and a truly comparative approach.

In what follows we shall label as *attention* the behavior under consideration, and mean by it any behavior, motor or perceptual, which has as its end-state contact between the organism and selected portions of its environment. Our major problem is to specify the conditions of organism-environment interaction that determine *what* gets selected. Those stimuli which are the objects of attention we label *goal stimuli*. We would distinguish *goal stimuli* from *means-end stimuli*: the latter represent the routes by which goal stimuli are attained. It is with respect to the selection of means-end stimuli that game theory and decision theory are concerned. Our primary concern is with the selection of goal stimuli.

### THE AROUSAL OF ATTENTION

In this section we shall discuss two experimental techniques that have been found successful in arousing attention. We shall argue that common to these distinct techniques is the production of *change*, whose relation to attention was early recognized by Pillsbury (17).

<sup>1</sup> Now at Yale University.

### *Temporal Change in Stimulation*

There are many experiments which utilize a temporal change in stimulation to arouse attention. Illustrative of these is an experiment by Dember (6). Rats on Trial 1 were allowed to see into the two arms of a T maze but were prevented by means of glass partitions from entering either arm. During the 10-minute exposure period one arm was black, the other white. The animals were removed from the maze, the partitions removed, and one arm was changed so that both arms were either white or black. The rats on Trial 2 were reintroduced into the maze and, faced with a choice between two arms equal in brightness, 17 out of 20 entered the arm whose brightness had been changed from Trial 1 to Trial 2. By the sign test, the above results are significant at the .001 level of confidence (one-tailed). Undergraduate students have repeated Dember's experiment several times since then with almost identical results.

Results analogous to these have been reported by Thompson and Solomon (18) and Berlyne (1, 2) with rats, and recently by Welker (20) with chimpanzees. It is to conditions like the above, involving operations that produce a temporal change in stimulation, that the term "novelty" most aptly refers.

It should be noted that a temporal change in stimulation arouses attention only if the change produces a discrepancy between what is observed by the subject and what is expected. The motivational implications of discrepancy have, of course, been emphasized by Hebb (9) and McClelland (12).

The experiments referred to above utilize change in stimulation over discrete periods of time. In one of Breese's many studies on binocular rivalry (4) a continuous change was effected in one of the two stimuli by means of a device which moved the stimulus over the field. The moving stimulus was found to dominate attention almost completely.

To summarize, an important experimental technique for arousing attention is to present the organism with stimulation which is discrepant from the organism's expectation. This may conveniently be accomplished either by exposing the organism to a stimulus and then changing the stimulus in some manner, or by producing a continuous change in the stimulus.

It is obvious that the indicator response, from which attention is inferred, may vary considerably, depending on the species studied and on the experimental conditions. We especially would emphasize that in the present analysis no distinction is made genotypically between "approach" and "avoidance" behavior. Hebb's chimpanzees, who find the isolated head of one of their species so discrepant as to be emotionally disturbing, nevertheless are certainly attending to it.

### *Spatial Change in Stimulation*

An analysis similar to that above may be applied to the case of goal stimuli whose ability to arouse attention is produced by spatial dishomogeneity. Breese (4), for example, found that one of the two fields in a binocular rivalry experiment could be made dominant simply by adding lines to it. Similar results are reviewed by Vernon (19). The experiments of Montgomery (15) are also pertinent here. His rats learned to choose the arm of a Y maze leading to a Dashiell maze within which they could wander, as opposed to the arm leading to a plain goal box. Welker's chimpanzees preferred the more spatially heterogeneous of a pair of stimuli (20).

Spatial heterogeneity is a characteristic to which the term "complexity" has been often applied. Complexity, like novelty, may be transformed into a psychological variable by defining it in terms of change in stimulation, in this case spatially induced.

While the distinction between the



temporal and spatial factors by which change is induced is useful in describing some of the experimental techniques available for producing goal stimuli, psychologically it is an unnecessary distinction. Spatial change implies some sort of scanning process on the part of the individual; the scanning of a spatially heterogeneous stimulus is equivalent to movement by the stimulus, over time, past the "stationary" individual. In this sense spatial heterogeneity has the same effect psychologically as temporal change of stimulation.

#### SYMBOLIZATION OF CHANGE

A useful set of symbols for abstracting the above ideas is offered in Coombs's measurement theory (5). The symbol  $Q_{hj}$  is the measure of a stimulus,  $j$ , on some attribute for an individual,  $i$ , at the moment,  $h$ . The symbol  $C_{hj}$ , analogously, is the measure of an individual,  $i$ , on some attribute for a stimulus,  $j$ , at the moment,  $h$ . The measures and attributes in this theory are not properties of the physical stimulus events or objects, but are defined by means of the discriminatory responses to such objects or events on the part of observers who are different from the experimenter. While the  $C$  values in Coombs's theory are ordinarily applied to such attributes as "arithmetic ability," "attitude toward the church," etc., we may generalize the concept of  $C$  value to include expectancy. Thus, an individual at time  $h$  may have a "brightness measure," not in the sense of intelligence or amount of light reflected from his face, but in the sense that he comes to a stimulus with an expectancy ( $C_{hj}$ ) about the brightness value ( $Q_{hj}$ ) of that stimulus. In the Dember (6) experiment, for example, the rat approaches the choice point on Trial 2 with a brightness measure for each arm,  $C_{hj}$  and  $C_{hk}$ . These measures correspond to the traces left by the two arms on the exposure trial.

A third measure derives from the interaction between  $C$  and  $Q$ :  $P_{hj} = |C_{hj} - Q_{hj}|$ .  $P_{hj}$  symbolizes the absolute value of the discrepancy between the expected and the observer;  $P_{hj}$  measures the novelty of a stimulus. In the example above, if  $j$  is the arm whose brightness was changed between trials, and  $k$  is the arm that remained the same, then it may reasonably be assumed that for the rat  $P_{hj} > P_{hk}$ .

Implied in this analysis is the assumption that, despite the brightness change, the maze arm,  $j$ , has retained its identity. It is the identity of  $j$  which makes meaningful the process implied in  $|C_{hj} - Q_{hj}|$ . This assumption, as Berlyne (1) points out, is made by McDougall (13) in his discussion of the curiosity instinct. The problem of stimulus identity has been emphasized recently by Kivy, Earl, and Walker (10), who demonstrated the importance of context in the production of satiation: exposure to a brightness equal to that of one of the arms of a T maze did not influence subsequent choice, if that exposure occurred other than in the T-maze choice point itself.

It is possible to apply to complexity the symbolization suggested above for novelty. Consider as an example of a complex stimulus a pattern of vertical black and white stripes. Suppose the subject scans the pattern from left to right, and the acceptance angle of his scan is one stripe width. First he sees a black area which has a certain  $Q$  value and which leaves the individual with a corresponding  $C$  value. He scans farther to the right and sees white. Now his  $C$  value, aroused by the preceding scan, will interact with the  $Q$  value of the present scan to yield a  $P$  value—if stimulus identity is preserved from first to second scan. It is as though the subject says to himself on the first scan, "This stimulus is black," and then as he scans farther "No, it's white."

The measure of complexity,  $P$ , is in



effect a measure of response variability; the interpretation of this measure is conceptually equivalent to the interpretation of "amount of information." For example, to the extent that discrepancies can arise as a stimulus is observed, there will be uncertainty, or information in the stimulus. In the spatial case, the amount of information will decrease as one scan yields a  $C$  value which equals the  $Q$  value of the succeeding scan, rather than the  $Q$  value of the immediate scan. A stimulus is redundant to the extent that it fails to offer the possibility of discrepancy.

Note that the loss of information, in the above sense, results from the modification of the relevant  $C$  values of the individual. Operationally, the stimulus is just as heterogeneous as before, but the individual through contact with it has changed. Unfortunately, information theory cannot solve the problem of the objective measurement of complexity, although it may contribute to an eventual solution.

It should now be clear that both temporal and spatial change are operations which are potentially novelty- or complexity-producing, but are not necessarily so. The amount of information arising from any experimental technique might be assessed a priori by methods such as Fitts *et al.* (7) have suggested. But the measure of complexity,  $P$ , arising from the application of that technique can only be recovered from the discriminatory responses of the subject under study. Under some conditions it is possible to make a good guess about the relative amounts of complexity in a set of stimuli, provided one has sufficient knowledge concerning the traits and past experiences of the subjects. Thus, it is quite likely that the striped stimulus above will be more complex for most subjects, at least upon first contact, than a spatially homogeneous stimulus. Similarly, the changed

arm,  $j$ , in the Dember (6) experiment should be more complex (or novel) for the animals than the unchanged arm,  $k$ . But it must be emphasized that these are merely hunches on the part of the experimenter. To make possible experimentation amenable to unequivocal a priori prediction, techniques are required that permit the assessment of stimulus complexity for the subject to be used. It is to this problem that some of Coombs's (5) recently developed measurement methods are perhaps applicable—at least where the subjects possess some degree of linguistic ability.

### A MOLAR ANALYSIS

We may consider the  $Q$ s,  $C$ s, and  $P$ s of the preceding discussion as elements in a more molar system whose units are sets of these elements. Thus, we might assign to a stimulus, on each of its attributes, an over-all complexity value arising from some interaction among the elementary within-stimulus discrepancies: symbolically,  $Q'_{hij} = f(P_1, P_2, \dots P_n)$ . In this sense we speak of a stimulus as having a complexity value,  $Q'_{hij}$ , on each of its attributes. Note that  $Q'_{hij}$  is a dimensionless quantity, as is pointed out by Miller (14) in his discussion of "amount of information."

The individual may also be assigned a complexity value,  $C'_{hij}$ , on each attribute, corresponding in a generalized sense to his "ability" on that attribute. Just as an individual at a given time has a certain ability with respect to arithmetic problems—some he can "pass," some he cannot—so, too, he has an ability with respect to paintings, music, literature.  $C'_{hij}$  is the complexity in a stimulus on some attribute which the individual at time  $h$  prefers.

Analogous to the molecular  $C$  values,  $C'$  changes with experience. Unlike the  $C$  values, however, the  $C'$  change is unidirectional: with experience,  $C'$  takes on increasing values. The individual

can only get *more able*.<sup>2</sup> Now the only stimuli which can change  $C'$  are those with  $Q'$  values greater than  $C'$ . The individual in effect is "paced" by stimuli. It is intuitively reasonable, however, that there be only a limited range of  $Q'$  values which are effective in changing  $C'$ . Some stimuli may be too complex to act as pacers. A person just learning French will not improve on being exposed to Proust or Baudelaire, though he may be able to evaluate their work as much more complex than that in his elementary text.

We would define as a *pacer* a stimulus with a  $Q'$  value in the acceptable range above  $C'$ . Now, a set of stimuli will be attention-arousing if that set contains a pacer. We postulate that under that condition, *the individual will apportion his attention among the stimuli in the set in proportion to their similarity to the pacer, with the modal amount of attention applied to the pacer.*

As the individual has experience with the set, his  $C'$  value will increase, until such time that the set no longer contains a pacer. At that time the set will no longer have the ability to change  $C'$ , and thus will have lost its attention-arousing property.

Given free choice, the individual will cease responding to the set. Forced to respond to it, he will experience boredom, and under certain circumstances a great deal of misery, as in the perceptual isolation experiments of Bexton, Heron, and Scott (3). Given no assistance from his receptors, he may resort to creating new stimuli—i.e., he may hallucinate, dramatically as in the isolation experiments, or less so, as in autokinetic, binocular rivalry, and reversible figures.

#### CONCLUDING COMMENTS

The ideas presented in this paper are neither novel nor complex, but we be-

lieve that some such scheme as we offer is necessary if we are to progress past the stage of merely demonstrating behavior apparently embarrassing to orthodox Hullians. Of course, by insisting that complexity is a psychological variable, assessable meaningfully only by psychological measurement techniques, we emphasize the difficulty of research in this area. But such a position is the only realistic one.

We further complicate the issue, and get closer to reality, by making attention a function not only of stimulus complexity but also of the individual's complexity—which itself changes with experience. We know nothing about the parameters of change, or how they vary across attributes. There is, however, the suggestion in an experiment by Krech and Calvin (11) that some aspect of change is related to verbal intelligence, which at least indicates that research on this problem is possible.

A most important source of research difficulty which we have allowed ourselves to face lies in the distinction previously mentioned between means-end stimuli and goal stimuli. Any research on goal stimuli must include techniques of insuring that it is goal stimuli that are being studied, or techniques of partialing out the means-end aspect of the stimuli. In working with human subjects, one must also find ways of eliminating ego involvement—of creating situations where the subject is truly free to behave in accordance with his evaluation of the stimuli as goal objects per se. A complete theory, of course, would specify the conditions for accomplishing this, but until that is available the experimenter's ingenuity must suffice.

One final point should be made concerning the concept of complexity. We have deliberately avoided defining complexity as an attribute, but rather have made it a dimensionless measure of a stimulus on any attribute. This enables us to take what seems the reasonable

<sup>2</sup> Regression of  $C'$  results from anxiety; the individual in this paper is always nonanxious.



position that a stimulus may have a different measure of complexity on each of its attributes; analogously an individual may have a complexity value on each attribute: on independent attributes these values may change independently. Thus, one's complexity for music may increase by the day, while one's complexity for literature remains, in the absence of the appropriate experience, at the level of Mickey Spillane.

### SUMMARY

We have classified exploration, manipulation, and curiosity in the category of attention. Two experimental techniques for arousing this behavior have been analyzed. We have argued that these two techniques have a common psychological basis, change or discrepancy, which may be symbolically represented in Coombsian terminology. The possible application of Coombs's measurement techniques to the experimental measurement of the amount of psychological discrepancy has been suggested. Finally, at a molar level, we have defined the conditions under which a set of stimuli will have attention-arousing value and under which this property is lost. Underlying the molar analysis is the idea that the dynamics of attention are based on the ability of stimuli to increase the psychological complexity of the individual who perceives them.

### REFERENCES

1. BERLYNE, D. E. Novelty and curiosity as determinants of exploratory behavior. *Brit. J. Psychol.*, 1950, 41, 68-80.
2. BERLYNE, D. E. The arousal and satiation of perceptual curiosity in the rat. *J. comp. physiol. Psychol.*, 1955, 48, 238-246.
3. BEXTON, W. H., HERON, W., & SCOTT, T. H. Effects of decreased variation in the sensory environment. *Canad. J. Psychol.*, 1954, 8, 70-76.
4. BREESE, B. B. On inhibition. *Psychol. Monogr.*, 1899, 3, No. 1 (Whole No. 11).
5. COOMBS, C. H. A theory of psychological scaling. *Engng Res. Bull.*, 1952, No. 34.
6. DEMBER, W. N. Response by the rat to environmental change. *J. comp. physiol. Psychol.*, 1956, 49, 93-95.
7. FITTS, P. M., WEINSTEIN, M., RAPPAPORT, M., ANDERSON, N., & LEONARD, J. A. Stimulus correlates of visual pattern recognition: a probability approach. *J. exp. Psychol.*, 1956, 51, 1-11.
8. HARLOW, H. F. Mice, monkeys, men, and motives. *Psychol. Rev.*, 1953, 60, 23-32.
9. HEBB, D. O. *The organization of behavior*. New York: Wiley, 1949.
10. KIVY, P. N., EARL, R. W., & WALKER, E. L. Stimulus context and satiation. *J. comp. physiol. Psychol.*, 1956, 49, 90-92.
11. KRECH, D., & CALVIN, A. Levels of perceptual organization and cognition. *J. abnorm. soc. Psychol.*, 1953, 48, 394-400.
12. MCCLELLAND, D. C., ATKINSON, J. W., CLARK, R. A., & LOWELL, E. L. *The achievement motive*. New York: Appleton-Century-Crofts, 1953.
13. McDUGALL, W. *An outline of psychology*. London: Methuen, 1923.
14. MILLER, G. A. The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychol. Rev.*, 1956, 63, 81-97.
15. MONTGOMERY, K. C. The role of the exploratory drive in learning. *J. comp. physiol. Psychol.*, 1954, 47, 60-64.
16. NISSEN, H. W. The nature of the drive as innate determinant of behavioral organization. In M. R. Jones (Ed.), *Nebraska symposium on motivation, 1954*. Lincoln: Univer. of Nebraska Press, 1954.
17. PILLSBURY, W. B. *Attention*. London: Swan Sonnenschein, 1908.
18. THOMPSON, W. R., & SOLOMON, L. M. Spontaneous pattern discrimination in the rat. *J. comp. physiol. Psychol.*, 1954, 47, 104-107.
19. VERNON, M. D. *A further study of visual perception*. Cambridge: Cambridge Univer. Press, 1954.
20. WELKER, W. I. Some determinants of play and exploration in chimpanzees. *J. comp. physiol. Psychol.*, 1956, 49, 84-89.

(Received March 14, 1956)



## CONDITIONS AFFECTING THE AMOUNT OF INFORMATION IN ABSOLUTE JUDGMENTS

EARL A. ALLUISI<sup>1</sup>

*Laboratory of Aviation Psychology, The Ohio State University*

Certain concepts of information theory (9, 15) have been used in measuring man's ability to make absolute judgments. It has been observed, for example, that as the amount of input information to the human is increased, the amount of information transmitted by the human (both in bits/stimulus) will increase at first and then level off at some asymptotic value (10). This asymptotic value is the maximum amount of information transmitted ( $\text{Max } H_i$ ) under the specific set of experimental conditions employed. When conditions are optimized for a specific task,  $\text{Max } H_i$  will attain its highest possible value; it may then be called the *channel capacity for  $x$* , where  $x$  is the specific psychological function or task studied (e.g., the visual discrimination of brightness, the auditory discrimination of loudness, etc.). When  $x$  is an absolute discrimination task, the anti-logarithm of the channel capacity may be treated as an estimate of the span of absolute judgments of that stimulus dimension, i.e., as an estimate of the number of categories of stimulation that can be discriminated absolutely with an arbitrarily small percentage of errors (6).

Spans of absolute judgments may differ for different stimuli, however, so that the channel capacity obtained with one

stimulus or task may be higher or lower than the value obtained with another stimulus or function. Man's *Channel Capacity* may then be defined as the highest of all the channel capacities for specific tasks; capitalization might even be used, as here, to differentiate this one "highest possible value independent of task" from the several "highest possible values for specific tasks."

These two concepts should not be confused. A recent review by Miller (10) was concerned with Channel Capacity (capitalized), whereas the present paper is concerned with channel capacity (uncapitalized). Specifically, the present paper is intended as a review of the available data indicating under what conditions of experimentation an estimate of channel capacity (rather than of  $\text{Max } H_i$ ) might be obtained.

In scope, this review will be limited to a consideration of data concerning the absolute judgments of stimuli lying along simple dimensions. First, auditory stimuli will be considered, then visual stimuli, and finally a brief section on individual differences will complete the paper.

### AUDITORY STIMULI

Pollack (12) had  $O$ s identify frequencies by assigning numbers to different tones. The tones differed in equal logarithmic steps in the range from 100 to 8000 cps. Knowledge of results was given  $O$  after each response; i.e., after each response  $O$  was told which of the tones had been presented. The amount of information transmitted by the av-

<sup>1</sup> The author wishes to express his appreciation to Drs. P. M. Fitts, L. Schipper, and J. A. Leonard, who kindly read and criticized a final draft of this paper. He would also like to acknowledge that much of the stimulation for the present review came from reading a paper on Channel Capacity (10) and a book on the applications of information theory in psychology (13, especially pp. 141-386).

erage  $O^2$  was found to increase linearly up to about 2 bits/stimulus, and then to approach an asymptote of about 2.5 bits/stimulus, as the number of alternative tones within the stated range was increased from 2 to 14 (i.e., as input information was increased from 1 to 3.8 bits/stimulus). This Max  $H_i$  of 2.5 bits/stimulus seems to imply that the span of absolute judgments of frequency differences in auditory stimuli is in the neighborhood of about five categories.

Pollack also found, however, that information transmission varied as a function of both the total range of frequencies used and the spacings used between stimuli. For eight tones, the mean amount of information transmitted with three closely spaced series was only 1.73 bits/stimulus, whereas it was 1.90 bits/stimulus with three widely spaced series. When the entire range from 100 to 8000 cps was used, the mean amount of information transmitted with eight tones was

2.00 bits/stimulus, but when only part of this range was used it dropped to 1.725 bits/stimulus. The actual values for the eight-tone data ranged from 1.6 to 2.1 bits/stimulus for the different conditions (12, p. 747, Fig. 5).

The finding that information transmission varies as a function of the spacing used between the stimuli has been corroborated by Garner (4). In a study that required Os to identify different intensities of a 1000-cps tone by assigning numbers to the different tones, he concluded that "... greater information transmission will occur when the stimuli are spaced according to a criterion of equal discriminability" (4, p. 238).

The finding that information transmission varies as a function of the range of stimulation used has been definitively corroborated by Schipper (14), whose experiment had two conditions. In one condition, the number of intensities of a 1000-cps tone judged by Os was varied from 2 through 4, 6, 8, and 10 tones with a fixed 5-db interval between stimuli. In the second condition, stimuli were selected at equal db intervals to divide a fixed 45-db range into 4, 6, 8, and 10 tones; this fixed range was equal to that used with 10 tones in the first condition. The amount of information transmitted by the average O in the first condition was found to vary from 0.47 to 1.29 bits/stimulus as the number of tones judged varied from 2 to 10. In the second condition, no significant differences in the amounts of information transmitted by the average O (about 1.26 bits/stimulus) were found for the different numbers of tones.

<sup>2</sup> There are two ways of computing an "average" of the information transmitted per O. Assume that there are  $m$  stimuli,  $n$  responses, and  $k$  Os. Then, (a) the amount of information transmitted by each of  $k$  Os could be computed separately from each of the  $k$ ,  $m \times n$  stimulus-response matrices. The arithmetic mean of these  $k$  values is referred to in this paper as the "amount of information transmitted by the average O." On the other hand, (b) the data of the  $k$  Os could be pooled into a single  $m \times n$  stimulus-response matrix from which a single "average" information value could be computed for all  $k$  Os. This value for pooled data is referred to in this note as the "average amount of information per O." These two specific phrases are used in the present paper only where the mode of computation is definitely known.

It might be noted that the latter value can never exceed the former for the same set of data. It may, however, be less than the former to the extent of the constant differences among Os. In fact, the difference between these two "average" values might be used as a quantitative estimate of individual differences, i.e., an estimate of the reduction in information transmission that occurs as a function of the constant differences among Os.

These results imply two things. First, there is an optimal spacing of stimuli that will maximize the measured Max  $H_i$  for a given stimulus dimension. There are theoretical grounds for postulating that when the stimuli are spaced according to a criterion of equal perceptual discriminability, the value of Max



$H_t$  obtained will be higher than with any other spacing (6), i.e., adjacent stimuli so spaced should represent equal confusion tendencies, so that the over-all tendency will be to minimize total errors and thereby maximize the information transmitted by reducing the total equivocation.

The second implication is that the asymptotic value of information-handling performance obtained with any given stimulus dimension is dependent upon the range of stimulation used in obtaining that  $\text{Max } H_t$ . The greater the stimulus range, the closer will  $\text{Max } H_t$  approach the channel capacity for the function being studied. This means, further, that even when optimal stimulus spacing is used, the inherent judgmental accuracy of  $O$  is independent of the number of stimulus categories used experimentally only insofar as a fixed range of stimulus variation is employed.

#### VISUAL STIMULI

The range of stimulus variation used experimentally has also been found to affect information transmission with visual stimuli. Eriksen and Hake (2) had  $O$ s judge the sizes of visually presented squares. A statistically significant increase in the amount of information transmitted by the average  $O$  was found to occur when the range of sizes was increased from 40 mm. sq. to 80 mm. sq.; the change amounted to about .2 bit/stimulus and represented an average change of about 10 per cent in the  $\text{Max } H_t$  estimated. The number of stimulus categories used in the study was also varied over 5, 11, and 21 different sizes within each range, and  $O$  used 5, 11, and 21 response categories with each of the different numbers of stimulus categories. A significant interaction was found between the number of stimulus and response categories. In general, the amount of information transmitted by the average  $O$  was lower for a given

range when the number of response categories was fewer than the number of stimulus categories.

In another study,  $O$ s interpolated visually between two scale markers defining a .16-in.-long horizontal interval. The number of pointer positions used within the interval as stimuli for  $O$  to judge varied over the values 5, 10, 20, and 50. Approximately equal amounts of information were transmitted with all numbers of pointer positions 10 or greater. From these data of Hake and Garner (7), a  $\text{Max } H_t$  of about 3.25 bits/stimulus could be estimated. Because the  $\text{Max } H_t$  appears to be a function of range, however, this estimate of nine or ten categories should not be taken as an invariant measure of the span of absolute judgments of interpolated visual position. Rather, it should probably be taken as a specific estimate applicable only to the length, or subtended visual angle, of the specific horizontal interval employed.

The effect of stimulus spacing has recently been investigated in a study of hue judgments. Conover (1) had  $O$ s judge the hues of colored papers, selected from the Munsell 50-hue maximum saturation series, by assigning numbers to the different hues. Two different experimental conditions were used, but in each condition the same range of stimulation (a complete color circle) was employed. In one condition, ten  $O$ s made absolute judgments of the 25 different stimulus categories obtained by taking every other hue in the Munsell series. Then, the 25 Munsell hues were scaled for equal discriminability, and 16 equally discriminable hues were selected from this scale. A different group of 21  $O$ s made absolute judgments of these 16 hues in the second condition of the experiment.

The amount of information transmitted by the average  $O$  was 3.486 bits/stimulus in the first condition, and



3.524 bits/stimulus in the second. The difference between these two values was not statistically significant. This, however, does not invalidate previous findings that spacing affects information transmission, for the scaling of the 25 stimulus hues had indicated that the stimulus categories were roughly of equal discriminability as originally selected from the Munsell series. That is to say, the equal discriminability scale formed with the 25 stimulus hues was approximately linear (see 1, p. 39, Fig. 5), and so the stimuli in both conditions of the experiment were spaced essentially with equal-discriminability intervals.

The estimated span of absolute judgments of hue in the Conover study (1) was about 11 categories, or 3.5 bits/stimulus. This is a little higher than the eight or nine categories (3.1 bits/stimulus) estimated from the data of Eriksen and Hake (3), but it is in close agreement with the estimated number of absolutely identifiable spectral hues (between 10 and 12 hues) reported by Halsey and Chapanis (8). The estimate would not apply, of course, to just a portion of the range of hues, but only to the full circle of hues.

Information transmission has been found to be affected by other variables in addition to range and spacing. Sidor-sky and Slivinske (11, pp. 30-37) had *O*s judge the inclination of a line of .1-in. length. The line could appear at any of the 72, 5° positions in the 360° of a circle. They found that when knowledge of results was given, the amount of information transmitted by the average *O* was 4.87 bits/stimulus (about 29 categories), but that when no knowledge of results was given it dropped to 4.66 bits/stimulus (about 25 categories). With four *O*s, the difference between these two means was statistically significant at the  $p < .01$  level of confidence. The presence or absence

of an outline circle around the inclination symbol did not affect information transmission (11, pp. 38-39), nor did variations in the line length of the inclination symbol from .1 through .3 in. (11, pp. 40-41).

Ellipse-axis ratios between a straight line (ratio of minor-to-major axis equal to zero) and a full circle (ratio of minor-to-major axis equal to unity) were judged, both with and without knowledge of results in another study (11, pp. 50-56). Twenty stimulus categories were used in equal axis-ratio steps of .053 between zero and unity; the major axis was  $\frac{7}{16}$  in. long in all cases. With the major axis horizontally oriented, the amount of information transmitted by the average *O* was 3.24 bits/stimulus with knowledge of results, and 3.13 bits/stimulus without knowledge of results. With the major axis oriented vertically and no knowledge of results, the amount of information transmitted by the average *O* was 3.18 bits/stimulus. The differences among these three means were not statistically significant at the  $p = .05$  level of confidence with five *O*s. Approximately nine categories are to be associated with each mean as the estimated span of absolute judgments of ellipse-axis ratios.<sup>8</sup>

<sup>8</sup> These studies in the absolute judgments of hue, inclination, and ellipse-axis ratio are unique in that they each cover a full possible range of stimulus variation. In a sense, then, the spans of absolute judgments estimated from these studies are not range-bound as are the values estimated from other studies. The spans for hue, inclination, and ellipticity appear to be in the neighborhood of 11, 29, and 9 categories, respectively.

For inclination, the value of 4.87 bits/stimulus exceeds the largest channel capacity reported by Miller (10, p. 86) for a unidimensional stimulus (3.9 bits/stimulus for positions in an interval). It might be argued, however, that inclination is not psychologically a unidimensional stimulus; two studies (11, pp. 30-37, 42-49) have indicated that judgmental accuracy is about the same in each of the four quadrants of the circle.

One other study is pertinent to this discussion of the effects of knowledge of results, spacing, and range conditions on the amount of information transmitted in making absolute judgments. Sidorsky (11, pp. 57-69) had *O*s make absolute judgments of the sizes of small circles of light within the range from  $\frac{3}{64}$  to  $\frac{17}{64}$ -in. diameters. When 15 categories were used with stimulus intervals of  $\frac{1}{64}$ -in. diameter, the amount of information transmitted by the average *O* was 2.66 bits/stimulus (about six categories) with knowledge of results, and 2.37 bits/stimulus (about five categories) without knowledge of results. For five *O*s, the difference between these two means was statistically significant at the  $p < .02$  level of confidence. Moreover, when eight categories were used over the same range with stimulus intervals of  $\frac{2}{64}$ -in. diameter, the amount of information transmitted by the average *O* without knowledge of results was 2.51 bits/stimulus (about six categories). When eight categories were used over half the range (from  $\frac{7}{64}$  to  $\frac{14}{64}$ -in. diameters) with stimulus intervals of  $\frac{1}{64}$ -in. diameter and no knowledge of results, the amount of information transmitted by the average *O* was 1.74 bits/stimulus (about three categories). The differences between this last mean and the other three means were all statistically significant at the  $p < .01$  level of confidence. Equal discriminability scales constructed for the four conditions of the study indicated that relative discriminability remained about the same over comparable ranges of the scale.

The results of this study of size judgments indicate that information transmission is influenced more by the range of stimulus variation than (a) by the number of stimulus categories, at least above some minimum, or (b) by the spacing between stimulus categories within wide limits. This further emphasizes that in order to obtain an es-

timate of channel capacity for any given stimulus dimension, the widest possible stimulus range must be used; otherwise, the obtained asymptotic value of information-handling capacity should be regarded as a  $\text{Max } H_i$  specific to the stimulus range used experimentally.

#### INDIVIDUAL DIFFERENCES

The effect of individual differences on information transmission may be illustrated in two ways. First, the information-transmission value computed from the pooled data of  $k$  *O*s<sup>4</sup> will be equal to, or less than, the arithmetic mean of the  $k$  information-transmission values computed from the unpooled data of  $k$  *O*s.<sup>5</sup> The difference between the two estimates of "average" information transmission is a measure of the gain in information transmission that occurs if the specific *O* transmitting the message is known. Illustrations of the different estimates obtained with these two methods of computation are shown in Pollack (12, p. 746, Fig. 3) and in Garner (5, p. 374, Fig. 1).

The second way in which individual differences in information transmission are illustrated is in the variation of the  $k$  values computed from the unpooled data for  $k$  *O*s. For example, the amount of information transmitted by ten *O*s making absolute judgments of 25 Munsell hues was found to vary from 3.185 to 4.215 bits/stimulus (9 to 19 categories); for 21 *O*s making absolute judgments of 16 hues with equal discriminability spacings, it varied from 3.117 to 4.000 bits/stimulus (9 to 16 categories) (1, p. 33, Table VI). Similarly, for 12 *O*s making absolute judgments of the intensity of a 1000-cps tone with ten stimuli within a 45-db range (5-db intervals

<sup>4</sup> The "average amount of information transmitted per *O*," as used in this paper.

<sup>5</sup> The "amount of information transmitted by the average *O*," as used in this paper.



between stimuli), the amount of information transmitted varied from 0.73 to 1.64 bits/stimulus (1.7 to 3.1 categories) (14, Appendix I).

The data of these studies indicate that measures of information transmission, like most psychological measures, vary from individual to individual. Therefore, estimates of the span of absolute judgments should only be based on data collected from adequate samples of individuals if those estimates are to be taken as representative of the population.

#### SUMMARY

This paper reviews the available data pertaining to the conditions of experimentation that appear to affect the span of absolute judgments. In addition to individual differences, the ability to make absolute discriminations among a set of stimuli appears also to be a function of the following four experimental conditions:

*Range of stimulus variation.* The amount of information transmitted with (or the estimated span of absolute judgments of) any given stimulus dimension is dependent upon the range of stimulus variation used experimentally. In general, the greater the stimulus range, the greater the amount of information transmitted in bits/stimulus. Range seems to be the largest over-all determinant of information transmission with a given stimulus dimension.

*Spacing of stimuli.* The maximum amount of information transmitted with a given stimulus range and a given number of stimulus and response categories may be reduced by use of nonoptimal spacings between stimulus categories.

*Number of response categories.* The number of response categories seems to interact with the number of stimulus categories used experimentally. In gen-

eral, when the number of response categories is fewer than the number of stimulus categories, the amount of information transmitted appears to be lower than when the number of response categories is equal to (or greater than) the number of stimulus categories.

*Knowledge of results.* The amount of information transmitted when knowledge of results is given to *O* after each response appears to be greater than when no knowledge of results is given.

In conclusion, man's channel capacity for any given task is defined as the highest possible asymptotic level of performance (highest  $\text{Max } H_i$ ) obtainable with that task. To measure channel capacity experimentally, optimal conditions of range, spacing, number of response categories, and knowledge of results would have to be used. Otherwise, the value obtained will be something less than the channel capacity—a  $\text{Max } H_i$  specific to the experimental conditions employed.

#### REFERENCES

1. CONOVER, D. M. The amount of information in the absolute judgment of Munsell hues. Unpublished doctor's dissertation, The Ohio State Univ., 1955.
2. ERIKSEN, C. W., & HAKE, H. W. Absolute judgments as a function of the stimulus range and the number of stimulus and response categories. *J. exp. Psychol.*, 1955, 49, 323-332.
3. ERIKSEN, C. W., & HAKE, H. W. Multidimensional stimulus differences and accuracy of discrimination. *J. exp. Psychol.*, 1955, 50, 153-160.
4. GARNER, W. R. An equal discriminability scale for loudness judgments. *J. exp. Psychol.*, 1952, 43, 232-238.
5. GARNER, W. R. An informational analysis of absolute judgments of loudness. *J. exp. Psychol.*, 1953, 46, 373-380.
6. GARNER, W. R., & HAKE, H. W. The amount of information in absolute judgments. *Psychol. Rev.*, 1951, 58, 446-459.
7. HAKE, H. W., & GARNER, W. R. The effect of presenting various numbers of discrete steps on scale reading accuracy. *J. exp. Psychol.*, 1951, 42, 358-366.



8. HALSEY, RITA M., & CHAPANIS, A. On the number of absolutely identifiable spectral hues. *J. opt. Soc. Amer.*, 1951, 41, 1057-1058.
9. MILLER, G. A. What is information measurement? *Amer. Psychologist*, 1953, 8, 3-12.
10. MILLER, G. A. The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychol. Rev.*, 1956, 63, 81-97.
11. MULLER, P. F., JR., SIDORSKY, R. C., SLIVINSKE, A. J., ALLUISI, E. A., & FITTS, P. M. The symbolic coding of information on cathode ray tubes and similar displays. *USAF, WADC Tech. Rep.*, 1955, No. 55-375.
12. POLLACK, I. The information of elementary auditory displays. *J. acoust. Soc. Amer.*, 1952, 24, 745-749.
13. QUASTLER, H. (Ed.). *Information theory in psychology*. Glencoe, Ill.: Free Press, 1955.
14. SCHIPPER, L. An analysis of information transmitted to human observers with auditory signals as a function of number of stimuli and stimulus intensity interval size. Unpublished doctor's dissertation, Univer. of Wisconsin, 1953.
15. SHANNON, C. E., & WEAVER, W. *The mathematical theory of communication*. Urbana: Univer. of Illinois Press, 1949.

(Received May 9, 1956)

Bureau Ednl. Psy. Research

DAVID HALL TRAINING CO.

Date

1956

## DISCRIMINATION LEARNING THEORY: UNIPROCESS VS. DUOPROCESS<sup>1</sup>

HARRY F. HARLOW AND LESLIE H. HICKS<sup>2</sup>

*University of Wisconsin*

A clear-cut differentiation may be made between uniprocess and duoprocess learning theories. Uniprocess theories assume that a single, basic physiological process, be it inhibition, excitation, or *X*, underlies habit formation and that a single interrelationship exists between this single process and reward, on the one hand, and nonreward on the other. The duoprocess learning theories assume that underlying learning there are two basic physiological processes, commonly called "excitation" and "inhibition," excitation being the resultant of reward, and inhibition the resultant of nonreward.

In this paper, interest is limited to the processes underlying discrimination learning. Many have advanced theories of discrimination learning, but, to date, Spence is the only writer to formulate a theory that generates quantitative predictions of the course of learning of a discrimination problem involving discriminanda with qualitative differences, the commonest type of laboratory discrimination learning problem. His theory clearly specifies two processes to account for the learning. Before dismissing other quantitative theories it might be pointed out that Hull (3), also a two-process theorist, limited his treatment of discrimination theory to situations in which the discriminanda vary quantitatively along a single stimulus dimension. Gulliksen and Wolfe (1) have proposed a theory broad enough to cover discriminanda having qualitative

differences, but application of the theory to qualitatively varying stimulus objects would require preliminary experimentation to scale the objects psychologically, an impractical procedure in the present situation.

Any duoprocess theory of discrimination learning must postulate that reward and nonreward have differential effects on learning. To account for the demonstrated phenomena of discrimination learning, it is probably also essential to duoprocess theory to assume, as did Spence (7) in his 1936 theory, that the differential effects are not constant throughout the course of learning but vary in some systematic manner. Spence postulated that reward leads to an increment in excitatory tendencies toward the stimulus components varying in amount per reinforcement during the learning of the problem so that the sequence of increments might form a parabolic curve. Thus, increments in excitatory tendency would be small early in learning, increase to a maximum in the middle course, and then decrease late in learning. Similarly, he postulated that nonreward results in a decrement in excitatory tendencies, regarded as "an active negative process, inhibition, which, adding itself in algebraic fashion to the positive excitatory tendencies, results in lowered strength values" (p. 430). He believes that successive nonreinforcements produce decrements or inhibitory tendencies increasing in amount as the strength of the response increases. Spence assumes that the successive decrements form a linear curve, although the actual form might be curvilinear. He was careful

<sup>1</sup> The research presented in this paper was supported by funds received from the Graduate School of the University of Wisconsin.

<sup>2</sup> Now at Howard University.

to point out that the exact forms of the curves of excitatory and inhibitory tendencies are not important to the theory, but it is implicit to the theory that the forms of these functions differ.

Unfortunately, in the two decades that have passed no one has apparently attempted a direct test of this crucial theoretical problem. It appears obvious that a number of tests favorable or unfavorable to the outlined duoprocess learning theory can be made, but there is some question whether or not the theory *per se* can be tested.

A direct measure of the relative or absolute increment or decrement produced on specified trials in an object discrimination problem is difficult or impossible because the patterns of the preceding rewards and nonrewards become so complex. Thus, no one has been able to assess the strengthening and weakening effects of reward and nonreward for Trials 7, 17, 27, and 37 or for any other selection or succession of trials. Furthermore, individual learning constants enter into the equations for predicting increments and decrements. Another difficulty in the theory outlined by Spence, as far as testing it is concerned, lies in the fact that the absolute or relative strengthening or weakening effect of reward or nonreward on any particular trial is influenced by original object and position preferences which the animal brings to the discrimination situation.

Spence's theory was designed, of course, to account for the learning of individual problems, but, because of the difficulties cited, a test of the theory in the single-problem situation is probably impossible. In a multiple-problem, or learning-set, situation, with a group of subjects, some of the difficulties can be controlled: object and position preferences can be balanced, and individual differences in the learning equation lose power through averaging. Likewise,

patterning of reward and nonreward trials can be controlled, so that the effects of reward and nonreward can be assessed at any stage of learning-set development. But the introduction of the multiple-problem technique introduces new variables whose possible influences might be a source of concern to those aligned with the Spence theory.

Sufficient data on learning-set development exist to justify, in view of the writers, the use of learning-set technique to test the operation of one vs. two processes in discrimination learning, and, moreover, they would hold that the results have relevance to the Spence theory of discrimination learning even though these results cannot be regarded as a pure test of the theory.

Object discrimination learning-set formation is conventionally described by plotting the percentage of correct Trial-2 responses of a series of problems. The resultant curve is typically either exponential or S shaped, depending upon the nature of the subjects, the pretest conditions, and the difficulty of the kind of object discrimination problem chosen. The performance increments resulting from the successive trials of a series of problems mirror, or can be made to mirror, increments resulting from successive trials of a single discrimination problem. If two processes operate in the learning of learning sets, differential performance curves should appear on Trial 2 of Problems 7, 17, 27, and 37 if the first trial of these problems was rewarded in the one instance and not rewarded in the other. If the successive Trial-1 responses of a set of problems are made correct or incorrect an equal number of times within succeeding small groups of problems, and if the learning of learning sets results from the two processes of "excitation" and "inhibition," the learning set curves for Trial 2 of this set of problems should be of one form for the problems that had



Trial 1 rewarded and of another form for the problems that had Trial 1 unrewarded. If uniprocess learning theory is correct, the two learning curves should have the same form throughout their course.

There is presumptive evidence that Trial-2 performance on successive discrimination problems is actually homologous with the successive-trial performance on a single discrimination problem. Meyer (4) has demonstrated that the scores made on Trial 2 of successive problems fall on the intraproblem learning curve, and that equal amounts of problem practice translate segments of the learning curve equal distances along a trial-unit abscissa. Although the Meyer measures were made on discrimination reversal problems, the results must hold for discrimination learning, since the only difference between these two problems for the monkey relates to the relatively unimportant stimulus-perseveration error-factor (2). Certainly the intimate interrelatedness of intraproblem and interproblem learning suggests a communality of basic mechanisms.

The actual problem which the present experiment is designed to test, however, is the operation of uniprocess vs. duoprocess learning in the acquisition of object discrimination learning sets.

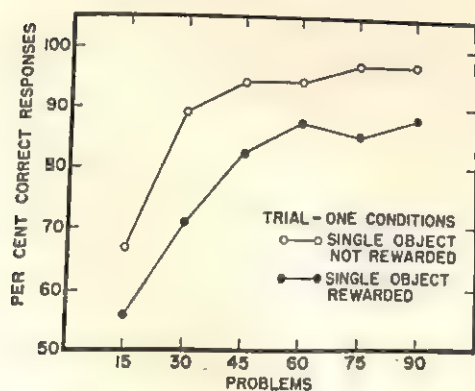


FIG. 1. Discrimination learning set curves based on Trial 2 responses.

## METHOD

Eight adolescent rhesus monkeys (No. 190, 191, 193, 195, 196, 197, 198, 199) served as subjects. They had been previously tamed, adapted to the test room and apparatus, and used in a latency experiment involving displacement of a single wooden block (5). The animals were tested in the Wisconsin General Test Apparatus with a two-hole tray. Stimuli were 360 pairs of objects differing multidimensionally.

The standard noncorrection trial procedure was utilized (2), with a raisin as the reward on correct trials. All problems were six trials in length, and 12 such problems were given each animal each day for 30 days. The position of correct and incorrect stimulus objects was controlled by use of balanced positional sequences (2).

The daily set of 12 problems for each S consisted of 3 problems of each of four kinds, differing only in first-trial conditions, as follows:

*Condition A.* A single object was presented over one of the two holes on the tray on Trial 1 and was not rewarded. On Trials 2 to 6 this object continued as the incorrect object paired with a second object which was consistently rewarded.

*Condition B.* A single object was presented over one of the two holes on the tray on Trial 1 and was rewarded. On Trials 2 to 6 this object continued as the rewarded object paired with a new object, consistently not rewarded.

*Condition C.* Two objects were presented on Trial 1, and response to either was not rewarded. The object which the monkey selected on Trial 1 was the incorrect object on Trials 2 to 6.

*Condition D.* Two objects were presented on Trial 1, and response to either was rewarded. The object which the monkey selected on Trial 1 was the correct object on Trials 2 to 6.

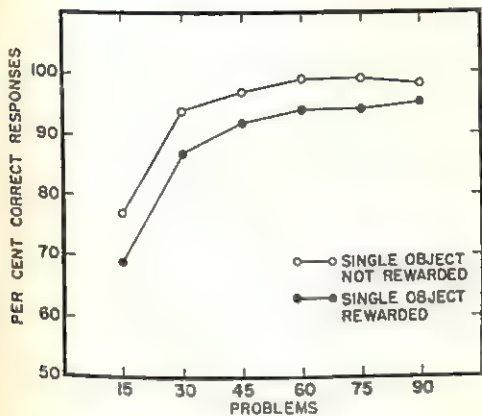


FIG. 2. Learning set curves based on Trial 2-6 responses.

Since Conditions A and B were the only conditions to balance stimulus and position preferences on Trial 1, they were regarded as the conditions critical for the test of uniprocess vs. duoprocess theory, and primary attention is directed to them.

### RESULTS AND DISCUSSION

The performance of the monkeys on the 90 successive Trial-2 responses obtained for each animal under the A and B conditions is presented in Fig. 1. There is no indication that the two curves differ in form or that they fail to parallel each other continuously. Furthermore, there is no indication that either curve changes its rate of gain relative to the other curve at any point.

The fact that performance is consistently better under Trial 1 nonreward than Trial 1 reward is immaterial, since it is differential curve form rather than a constant rate difference which would be expected if reward and nonreward showed differential strengthening and weakening effects during the course of learning.

The obtained constant rate differences had been expected, empirically, in accordance with the findings of Moss and Harlow (6), and also theoretically, since nonreward minimizes response

shift and reward maximizes it (2). Such data as exist (2) would indicate that the response-shift error-factor for Condition-B problems would either not be materially reduced during the course of the experiment or would be reduced only near the end. Thus, it would either not affect the form of the curve, or it would affect it only near the end of the experiment, in this case reflected as a gradual rise of the Trial-2 Condition-B curve as compared with the Trial-2 Condition-A curve.

The performances of the monkeys on the 90 successive mean Trial 2 to 6 responses made under Conditions A and B are presented in Fig. 2. Again, the two curves appear to be highly similar, if not identical, in form. It is possible that at the last plotted point the Condition-B reward curve is accelerating relative to the Condition-A nonreward curve, but the differences are well within the range of experimental error, and would not be unexpected in terms of the described operation of the response-shift error-factor (2).

The learning curves for all four training conditions during Trials 2 to 6 are depicted in Fig. 3, even though it was considered that Conditions C and D are not critical tests because of lack of control over stimulus preferences. If neither object is rewarded on Trial 1, the mon-

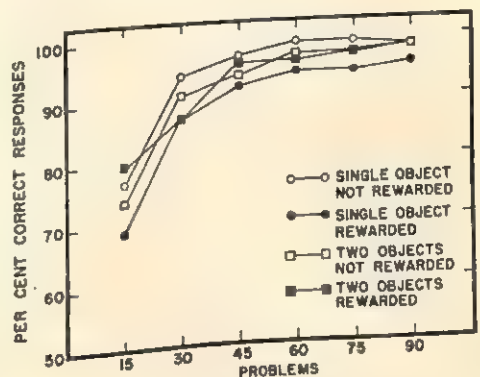


FIG. 3. Learning curves for the four conditions of first-trial reward.

key's preferred object, as indicated by Trial-1 choice, becomes the negative object, and the animal must choose the nonpreferred object on subsequent trials. If both objects are rewarded on Trial 1, the preferred object is the correct one for the trials that follow. The high performance under Condition D at the initial plotted point, and at no other point, was predicted from the fact that stimulus preference is an important factor influencing choice early in learning, and then decreases sharply in importance after the early stages of object discrimination learning-set formation by rhesus monkeys (2). It may be noted that the composite curves for the A and B, single-object, conditions, and for the C and D, double-object, conditions are almost identical at all points after Problems 1 to 15.

It is quite obvious that the present experiment is not a direct test of Spence's discrimination learning theory, because the present experiment measures learning in the learning-set situation. The data appear, however, unequivocally to support a uniprocess interpretation of learning, because reward and nonreward operated in a parallel and constant manner throughout the entire course of the experiment in influencing learned performance.

No single experiment can possibly settle a theoretical problem as crucial as that involved in choosing between duoprocess and uniprocess learning mechanisms. The present experiment does present evidence in favor of the uniprocess conception of learning, and, more important, it suggests the possibility of other direct attack on the problem.

If a uniprocess theory is correct, it must survive other experimental tests. For example, if we accept that learning is ascribable to a single process, we must assume that the generalization curves to the rewarded and to the unrewarded

training stimuli will be identical, or that any departure from identity can be explained by the operation of independently defined mechanisms, such as response shift. Similarly, the temporal decay which should result from increasing the intervals of delay between successive trials of discrimination learning problems must be constant for rewarded and unrewarded stimuli, or must be attributable to independently defined mechanisms. Specifically, if we measure the learning on Trial 1 in terms of performance on Trial 2, and introduce four intertrial delay intervals such as 10, 20, 30, and 40 seconds, the delay functions to initially rewarded or unrewarded stimuli should be identical in form. Similar results should also obtain in delayed-response testing.

On the basis of preliminary studies we know that adequate testing of uniprocess theory in these described situations is difficult because of the rigid controls which must be established over past experience, stimulus preference, and the operation of the powerful response-shift error-factor. Our preliminary studies consistently show greater learning to the unrewarded stimulus in all situations; the results are indeterminate in regard to the form of the generalization curves or the temporal decay functions. The results are, however, of such a nature as to suggest that definitive tests can be made by improved precision of testing and better understanding of the factors which operate to produce error in the nonspatial discrimination learning situation.

If a uniprocess learning theory is accepted, it is interesting to speculate concerning the fundamental nature of this mechanism. There are three possibilities: the mechanism may be excitatory, inhibitory, or undefined—a mechanism *X*. Wisdom or cowardice would dictate mechanism *X*. Hebb apparently prefers excitation, and the authors prefer



inhibition. At the present time there appears to be no direct test among the three mechanisms, and, this being true, choice is dictated merely in terms of which thesis best integrates existing learning data. Analysis of the nature of errors during learning and a review of the phylogenetic literature on learning have influenced the authors in their choice. The problems and perils of inhibitory learning theory are beyond the scope of this paper, but the assumption of a position offers new approaches in visualizing learning and the development of new families of problems.

## REFERENCES

1. GULLIKSEN, H., & WOLFLE, D. L. A theory of learning and transfer: I. *Psychometrika*, 1938, 3, 127-149.
2. HARLOW, H. F. Analysis of discrimination learning by monkeys. *J. exp. Psychol.*, 1950, 40, 26-39.
3. HULL, C. L. *A behavior system*. New Haven: Yale Univer. Press, 1952.
4. MEYER, D. R. Food deprivation and discrimination reversal learning by monkeys. *J. exp. Psychol.*, 1951, 41, 10-16.
5. MICHELS, K. M. The acquisition and retention of single stimulus responses by monkeys as a function of fixed-ratio reinforcement. Unpublished doctor's dissertation, Univer. of Wisconsin, 1953.
6. MOSS, EILEEN M., & HARLOW, H. F. The role of reward in discrimination learning in monkeys. *J. comp. Psychol.*, 1947, 40, 333-342.
7. SPENCE, K. W. The nature of discrimination learning in animals. *Psychol. Rev.*, 1936, 43, 427-449.

(Received April 3, 1956)

## CREATIVE ABILITIES IN THE ARTS<sup>1</sup>

J. P. GUILFORD

*University of Southern California*

In 1950, the writer set forth some hypotheses concerning the component abilities that were believed to be needed to account for creativity (1). These hypotheses were developed by way of preparation for systematic studies of this phenomenon by a combination of experimental and factor-analytic approaches. In order to complete the setting of these studies, which is important in preparation for some of the things to follow, something must be said about the scope of the studies of creativity.

The aptitudes project<sup>2</sup> has not been confined to the study of creativity, but has investigated all types of thinking abilities, including also those traditionally known as reasoning abilities, and those we chose to include under the headings of planning and evaluation. This inclusiveness was fortunate, for we find that the whole area of thinking abilities or functions is rich with interrelations and parallels. The understanding of some parts of this total area is very helpful in understanding others.

On the other hand, the studies of creativity proper, up to the present, have been more limited than they might have been. In setting up hypotheses concerning the component abilities in creativity, we were guided mostly by the kinds of creative activity recognized as such in scientists, engineers, inventors, and in supervisory and administrative personnel; in other words, types of per-

sonnel that are of concern in the military setting. We did tolerate the general hypothesis that the abilities that make these kinds of personnel creative might be the same as those that make the painter, the composer, the writer, and others creative, but we did not reject the contrary hypothesis, for we had no basis for doing so. To be sure, it would be a simpler outcome to find that the same qualities of fluency, flexibility, and originality, for example, account for performances of artists and scientists alike. But in our research we have never been very strongly influenced by the goal of simplicity. We have seen that all too often the compulsion of this goal has been unfortunately restricting of the investigator's outlook.

We did favor the notion that creativity, whatever its range of application, is by no means a unity but is rather a collection of different component abilities or other traits. Our results have definitely supported this general point of view. They also suggest the hypothesis that in the areas of the performances of the graphic artist and the composer, at least, we shall find new factors; factors that are distinct from those that are important in creativity of scientists, technologists, and managers, yet that are parallel to them. It is from the information concerning abilities that we have already investigated that we can deduce something about creative abilities that we should find to be important in the arts. The artistic, creative abilities that I shall mention are thus mainly hypothetical, but I should say that there are better precedents for these hypotheses than for those presented in 1950. We have made a beginning toward re-

<sup>1</sup> Based upon a paper presented in the symposium on "Aspects of Creativity" at the convention of the American Psychological Association, September 4, 1956, in Chicago, Illinois.

<sup>2</sup> Under Contract N6onr-23810 with the Office of Naval Research, monitored by the Personnel and Training Branch.

lating some of the known factors to the art of writing. It is hoped that the presentation of the hypotheses in this article will stimulate other investigators to test them in connection with other arts as well.

In general, the support for the expected factors thought to be important in the arts lies in the systematic nature of the whole collection of thinking factors, also the memory factors. Enough of the thinking and memory factors are known for us to see the lines of a system. A conception of the entire collection of intellectual factors has been presented in a recent issue of the *Psychological Bulletin* (2). It will pay us to review briefly the features of that system that are relevant in support of factors that are predicted to be important in the arts.

#### THE SYSTEM OF INTELLECTUAL FACTORS

One of the significant principles of the system is that the factors fall into three parallel groups, depending upon the kind of material involved in the activity. Let us think of them as being in three parallel columns. Psychologists have had a long-standing recognition that different abilities are involved in verbal tests on the one hand and nonverbal tests on the other. In nonverbal tests the psychological material dealt with by the examinee is in the form of figures, letters, numbers, or other symbols. Our project results show that we must make a further distinction within the nonverbal area. The consequence is that the intellectual factors tend to come in groups of three parallel abilities or traits. For example, there are three abilities for seeing relationships between things. One of them applies to relations between perceived figures. A second has to do with relations between meanings or concepts. There is a third relationship factor that has to do with the ability to see relationships between such materials as letters, numbers, or

other simple symbols. In the latter case, it is neither their figural nor their meaning properties that determine the relationships; it is some other property. We have called this category of factors the "structural" group. In everyday life, the structural type of thinking is perhaps most evident in mathematics. It does not appear that the structural factors have much significance for the arts, as such, and that we shall have to look for the significant artistic abilities among the figural and conceptual factors.

One or two additional comments will help to show just where the artistic-creative factors fit into the general scheme of intellect. The thinking factors seem to fall into three general groups of another kind, in a cross classification. This grouping is based upon the kind of *action* performed. There is a group of *cognition* factors, a group of *production* factors, and a group of *evaluation* factors. We become aware of the things with which we are confronted; we produce something of our own in response to that awareness, or something that it calls for; and we evaluate our products of thought. A total creative act involves all three aspects—cognition, production, and evaluation. A schematic view of all the classes of intellectual factors is shown in Figure 1.

In view of the active nature of creative performances, the production aspects or steps are most conspicuous and

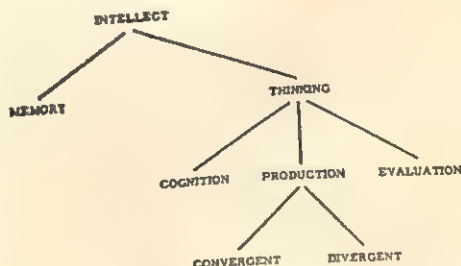


FIG. 1. Diagram of the major categories of intellectual factors and their logical relationships.



probably most crucial. Among the productive-thinking abilities another logical distinction appears. With some productive-thinking factors, and the tests that measure them, thinking must at some time converge toward one right answer; the significant type of thinking involved has been called "convergent" thinking. With other productive-thinking factors and their tests, thinking need not come out with a unique answer; in fact, going off in different directions contributes to a better score in such tests. This type of thinking and these factors come under the heading of "divergent" thinking. It is in divergent thinking that we find the most obvious indications of creativity.

This does not mean that convergent thinking and divergent thinking never occur together. They frequently do, in a total act of problem solving. Creative steps are necessary in solving new problems. Actually, we can hardly say there is a problem unless the situation presents the necessity for new production of some kind. Factors are abstractions of components from total activities. Some of the components are recognized as being more creative than others, for example, qualities of fluency, flexibility, and originality. These come under the category of the divergent-thinking aspect. While they may contribute toward reaching one right answer where that is demanded, they are more obvious in activities where this is not the case. In the arts there is usually *no* one right answer. Some answers are regarded as merely better than others. There is a matter of evaluation.

### FLUENCY FACTORS

Let us consider first the potential factors in the more obvious creative areas of fluency, flexibility, and originality. Our project results thus far have clearly indicated four fluency factors, two flexibility factors, and one original-

TABLE 1  
A TABLE OF THE PRODUCTIVE-THINKING  
FACTORS OF THE DIVERGENT-  
THINKING TYPE

Type of Result Produced	Type of Content		
	Figural	Structural	Conceptual
Words		<i>Word fluency</i>	<i>Associational fluency</i>
Ideas			<i>Ideational fluency</i>
Expressions			<i>Expressional fluency</i>
Shifts	<i>Flexibility of closure</i>	<i>Adaptive flexibility</i>	<i>Spontaneous flexibility</i>
Novel responses			<i>Originality</i>
Details	<i>Elaboration*</i>		<i>Elaboration*</i>

\* At present regarded as the same factor, but future results may indicate two separate factors.

ity factor. For the most part, our tests of these factors fall into the conceptual column. Table 1 is given to show the whole matrix of productive-thinking factors and their interrelationships.

Let us consider the fluency factors first. Two of them, *word fluency* and *associational fluency*, have to do with the production of single words. Tests of *word fluency* are best characterized by the fact that the words produced must meet specified structural requirements, such as listing words beginning with a certain letter or words ending in a certain suffix. Meanings or concepts are of no importance. The *word-fluency* factor thus falls in the structural column. Parallel to it, the *associational-fluency* factor is measured by tests that involve listing words having some meaningful requirement, such as listing synonyms or opposites for a stimulus word. To complete this triad of factors, there should be one involving the production of letter combinations that satisfy certain *figural* requirements, such as the activity of producing monograms or other artistic effects with words.

The factor of *ideational fluency* stands alone at present in another incomplete

triad. It is the ability to produce rapidly a succession of ideas meeting certain meaningful requirements. The number of words produced in each response may be one or several. For example, tests of this factor may call for the listing of things round, of ideas about a man going up a ladder, of titles for a story plot, or of predictions of consequences of events. Quantity is important but quality is not.

A parallel factor in the figural column would be an ability to produce a variety of artistic ideas in limited time. Rough ideas for themes, rough sketches, and the like would be sufficient output so far as this ability is concerned. At this point we must face a question that has general significance, beyond the *ideational-fluency* areas. This is the question whether ability to produce numerous ideas in the graphic arts is the same as the ability to do so in music.

There is a precedent at one place in the system of intellectual factors for a distinction between visual and auditory functions. This occurs among the memory factors. We have an apparent triad of memory factors in all of which the learning and retention of associative connections between contents is the important thing, and a second triad in which memory for the contents themselves is essential. In the latter triad, there is an ability to remember the substance of meaningful verbal material. Parallel to it are two factors, rather than one, having to do with remembering substance in figural form. There is a factor of *visual memory*, and this is separate from a factor of *auditory memory*. The latter involves memory for such things as melodies and rhythms. We may regard melodies and rhythms as auditory figures.

The separation of two figural-memory factors, visual and auditory, suggests that a similar distinction may be found elsewhere in the system, and hence pos-

sibly in the area of the production of figural ideas. This would mean that the ability to produce ideas in the graphic arts is distinct from the ability to produce ideas in music. With this separation of factors, we should have four ideational-fluency factors, not forgetting the one in the structural column, rather than the one factor already known. There could even be a fifth ideational-fluency factor connected with the kinesthetic sense, also in the figural column. This ability would presumably be of importance to the successful choreographer or creative acrobat.

The fourth known fluency factor is called *expressional fluency*. Thus far confined to verbal tests, this factor is recognized as an ability to put ideas into words. Tests requiring the putting of words together in appropriate, connected discourse are best measures of the factor. The distinction of this factor from *ideational fluency* is support for the common observation that it is one thing to have an idea and it is something else to be able to put it into words.

Three of the known fluency factors should go a long way to account for talent for writing. *Ideational fluency* should give the writer something to write about; *expressional fluency* should enable him to put it into appropriate words; and *associational fluency* should help to find words with the right shadings of meanings without the help of word-finding aids.

Is it likely that there are factors to complete a triad of expressional factors? The concept of expression is surely not foreign to the arts. Having a graphic or a musical idea is short of the total creative production. Putting the idea into appropriate organizations of figural material would be necessary to complete the process. And possibly, again, we shall find that expression in graphic form depends upon a different ability

than expression in musical form, just as both differ from expression in verbal form.

A general conception of creativity that calls for so many distinctions and separations of function may be somewhat surprising to readers. Why, after all, should there not be much in common between having ideas in the graphic arts, in music, and in writing? Why should there not be much in common in expressing ideas in the different media?

Notice that I have referred to such abilities as being distinct; not necessarily as being independent or uncorrelated. I suspect that there *is* something in common among parallel factors. This should not preclude their statistical and experimental separation, provided that performances of these different kinds or with different materials are not perfectly correlated when correction is made for unreliability. I suggest that we proceed to find out whether the factors are statistically separable and, if so, whether they separate along the line hypothesized or along some other lines. Then we can more appropriately raise the question about their interrelationships.

There is some evidence from everyday observation to lend support to the separateness of the expressive abilities. For example, the production of a popular song often involves the collaboration of a composer and a lyricist. To be sure, some individuals do both successfully. Actually, the correlation could be zero between these two performances, and yet there would be by chance *some* individuals who show high status in both respects. The production of a motion picture, in which musical, graphic, and conceptual ideas are commonly expressed in blended combination, is a synthetic task of specialists, just as, more and more, even single cinema characters rep-

resent synthetic blends of talents of different performers.

#### FLEXIBILITY FACTORS

The two flexibility factors that we have found differ in more than one respect. One difference is that one factor is found in verbal tests and the other mostly in nonverbal tests. One therefore belongs in the conceptual column and the other in the structural column, or possibly it cuts across structural and figural columns; we are not sure. The other difference is in the role that each factor plays or the degree of compulsion vs. freedom involved.

The conceptual-flexibility factor is called *spontaneous flexibility* because the examinee shows flexibility on his own initiative; the test items do not necessarily require it. It is possible that this quality is a temperamental trait or a motivational trait rather than an ability; a disposition to avoid repeating oneself, or an urge to vary one's behavior. If this is true, the trait might be accounted for under the Hullian concept of "reactive inhibition," or under the concept of a general psychological refractory phase, or under the concept of satiation. Being quite general in its determination of behavior, such a trait might serve as the basis for very fanciful, creative imagination wherever it is found, for example, in artist and scientist alike.

The other factor in this area we have called *adaptive flexibility*, because it is important in the solution of problems—particularly those that require the discarding of familiar or habitual methods and striking out in new and unusual directions. We have more recently expected to find three factors of this kind, but thus far have not found them, probably because our test variations have been inadequate to effect their separation. To the extent that there are prob-



lems involved in the arts, this kind of ability or trait would seem to play a significant role. It remains to be seen whether an adaptive-flexibility factor that is unique to figural material is required.

#### ORIGINALITY

The one factor of *originality* seems to be rather general, in one sense at least. That is, it is indicated by varied tests; tests that require unusual or uncommon responses, remote associations or connections, or clever responses. The use of an unusual variety of tests has provided much opportunity for a separation into two or more originality factors along the lines of such differences. All the tests have been verbal, or have involved verbal meanings in some way. We have as yet provided no opportunity for finding a triad of originality factors, distinguished along the lines of the materials involved.

There is a possibility that the factor of *originality* will prove to be fundamentally a temperamental or motivational variable. For example, it might be a general set to be unconventional or to avoid repeating what other individuals do. A single trait of this kind could be expected to cut across material categories. There would then be one originality factor; not a triad. Some of our future research will be directed along these lines with regard to originality as well as with regard to flexibility.

We have already made a beginning toward relating fluency, flexibility, and originality factors to temperamental and motivational variables. At present, it does not appear that any of them can be accounted for on the basis of such nonaptitude variables as we are exploring in this connection. This leaves the way clear for testing the hypothesis that there are complete triads of factors in these three areas.

#### OTHER FACTORS RELATED TO CREATIVITY

In 1950, in addition to factors of fluency, flexibility, and originality, it was hypothesized that there would be an ability to see problems, an ability to analyze, an ability to synthesize, and an ability to redefine or reorganize objects of thought. The hypotheses concerning analyzing ability and synthesizing ability were rather decisively discredited by the results. In spite of the opportunities for such unitary abilities to make themselves known, they failed to appear. This does not mean that, in thinking, no such activities as analyzing and synthesizing take place, for too many activities can be described as such. The unitary abilities that individuals have in common and that have a bearing upon success, however, are better described otherwise. This kind of conclusion is not unique. No one would deny that we indulge in activities that are properly called thinking, and yet there is no generalized unitary ability to think. There are many thinking abilities, as previous discussion has demonstrated. An all-too-common error in psychology has been to assume that because a range of phenomena can be subsumed under a single name there is therefore a unitary function. Every such assumption must be tested by empirical procedures.

A factor was found that could be defined as the ability to see problems. It is a cognition factor rather than a production factor. It proved to be much less general than was originally expected, being confined to seeing defects and deficiencies in such practical matters as everyday gadgets and implements and in social institutions and practices. The tests that measure the factor have been exclusively verbal. Will the use of comparable nonverbal tests give us completion of a triad of problem-seeing factors? This hypothesis should be tested.

To the extent that the artist has problems, we may suppose that there are individual differences in ability to recognize them. The problem might be in the form of the need for a theme or a particular kind of theme, or in the form of expression or treatment, or in the use of techniques and implements. Among these would be problems involving figural properties of things. The triad hypothesis would lead us to expect little correlation between the ability to detect *such* problems and the ability to detect the kind represented in verbal tests.

The factor of redefinition involves the ability to desert one interpretation or conception of use of an object, or part of an object, and to adapt it to some new function or use. For example, the cover glass of a watch can be removed and used as a condensing lens to start a fire. How readily can the individual arrive at such a transformation? How good is he at improvising in similar situations in general? This variable is a divergent-thinking factor that involves the production of a shift of meaning of an object. Are there parallel factors involved anywhere in the arts?

Actually, there is a factor of *visualization*, which seems to be to the figural column what *redefinition* is to the conceptual column. The factor of *visualization* is the ability to think of changes or transformations of a figural kind in visually perceived objects, or in objects visually thought of. The relation of such an ability to work in the visual arts can be readily imagined. There might even be such a factor in the auditory field, enabling a composer or arranger to produce variations on a theme with changes in use of phrases so radical that they take on new values or functions.

A factor of evaluative ability was hypothesized, not as a contributor to the production of creative results but as a means of determining whether such re-

sults are good, suitable, correct, or adequate. In our investigation of this area of thinking, we gave ample opportunity for more than one evaluative-ability factor to emerge. There are different bases or criteria by which a product is judged. One is its logical consistency with known facts. Another is its less-than-logical consistency with other experiences. There are also different kinds of products to be judged, depending upon the kind of materials involved. We included tests with both figural and verbal material. At the time of the study of evaluative abilities, the third category of structural materials had not yet been recognized.

We found three general evaluation factors. *Logical evaluation* is an ability to judge products on the basis of their logical consistency with given facts. A factor called *experiential evaluation* seemed to fit the picture of an ability to judge products in terms of consistency with past experiences. In the interpretation of this factor, if the emphasis is placed upon ability to make use of past experiences in the act of judging, it could be a rather general ability. If, however, emphasis is placed upon the past experiences, we face the real possibility of many common factors of this kind, depending upon the more or less coherent bodies of information that people acquire, for example, mechanical, mathematical, and so on. As for the rest, the use of experience would be a rather specific matter.

A third factor, which was called *perceptual evaluation*, is of uncertain generality. It can readily be hypothesized that there are as many perceptual-evaluation factors as there are coherent areas of perceptual functioning. The variety of psychophysical judgments is, of course, almost unlimited. The tests that defined our *perceptual-evaluation* factor emphasized comparisons of lengths of lines and total sizes of figures. The factor we found may therefore have been



the more limited *length-estimation* factor that was previously known.

The whole area of evaluative abilities is still largely unexplored. I have hinted that we may expect to find a very large number of rather narrow evaluative-ability factors. As for evaluation in the arts, presumably the *logical-evaluation* factor would not apply. Experimental-evaluation abilities might account for aesthetic tastes in terms of aesthetic values. Perceptual-evaluation abilities would have much bearing on the acceptability of art forms, visual, auditory, or kinesthetic. They would perhaps be numerous and also generally of narrow scope.

The factors mentioned thus far are those we originally regarded as belonging in the creative category. Recognizing that some aspects of planning are also creative, certain newly obtained factors in that area could also be regarded as creative. But as the system of the intellectual factors developed, cutting across our original categories of reasoning, creativity, and planning, these category concepts have shrunk in importance. Furthermore, it became more apparent that, in the creative activity of everyday life, other abilities than those regarded as primarily creative also play roles to some degree. For example, is it not likely that a large vocabulary is desirable for the creative writer? Should not the developer of ideas in descriptive geometry be able to think readily in terms of visual-spatial arrangements? These two examples imply the usefulness of the factors known as *verbal comprehension* and *spatial orientation*, respectively. Norman C. Meier has also emphasized the finding that individuals with recognized artistic talents are unusually able to observe and to remember clearly things they perceive (3). This implies a high degree of the factor known as *visual memory*, an ability to remember visual content. The factor of

*auditory memory* may play a similar role for the composer.

Thus, a great number of primary mental abilities that would not be regarded as creative abilities nevertheless play their roles at times in creative work. We might say that minimal levels of such abilities are desirable, if not necessary, for success in various artistic activities. We might say that to that extent these are necessary but not sufficient conditions for creative production. The factors of fluency, flexibility, and originality, and the like, are not only necessary but, when possessed in adequate amounts, are sufficient. All of this, of course, assumes adequate motivating conditions, also. In the process of surveying the resources of creative artists of any kind, therefore, whether this is for the sake of better understanding of talent or for the practical purposes of prediction and guidance, it would be well to ask whether any of the intellectual factors may play a significant role, and where and how, if so.

#### SUMMARY

1. It is hypothesized that creative artistic talent is not a unitary or uniform commodity but is to be accounted for in terms of a large number of factors or primary mental abilities. From what is already known, we should expect that the creative abilities of artists will be found to involve some factors other than those among creative abilities in fields such as science and management.

2. Of the known factors, certain ones, of fluency, flexibility, and originality, are the most obviously creative abilities. All of them come under a general class of factors known as productive-thinking abilities and in a subclass of divergent-thinking abilities.

3. A developing system of all the intellectual factors indicates the relationships of the more creative factors to one another and to other factors. From



certain relationships and parallels, unknown factors that are probably important in the arts can be hypothesized with some confidence.

4. A full account of complete creative-artistic performance involves evaluative abilities and abilities that are not primarily creative, many of which are already known.

## REFERENCES

1. GUILFORD, J. P. Creativity. *Amer. Psychologist*, 1950, 5, 444-454.
2. GUILFORD, J. P. The structure of intellect. *Psychol. Bull.*, 1956, 53, 267-293.
3. MEIER, N. C. Factors in artistic aptitude: final summary of a ten-year study of a special ability. *Psychol. Monogr.*, 1939, 51, 140-158.

(Received October 3, 1956)

## A GRAPHICAL DESCRIPTION OF ROTE LEARNING

BENTON J. UNDERWOOD

*Northwestern University*<sup>1</sup>

It has often been pointed out that a learning curve for a group of *Ss* does not adequately reflect the fluctuations in the performance of the individual *S*. A curve representing the combined performance of a group of *Ss* usually shows a smooth increment in performance as a function of trials or time, whereas a curve for a single *S* shows increments, decrements, and plateaus. However, since these fluctuations occur at different points in time for different *Ss*, they apparently average out to show a smooth progression for the group. A group curve could reflect these fluctuations if (a) a given fluctuation is lawfully related to earlier performance of an *S*, and if (b) common reference points for all *Ss* could be used to measure this lawfulness. Hayes (1) has shown how this descriptive problem might be handled with data derived from the discrimination learning of rats. The present paper shows one way in which a group curve will reflect the individual fluctuations for data from rote learning. The method, with exemplary data, will be given first and then certain limitations and implications of the description will be pointed out.

Consider the situation when *S* is given a serial or paired-associate list and is asked to learn this list to a criterion of one perfect trial. Different *Ss* will require different numbers of trials to attain this criterion of mastery. Two general ways have been used to show the group learning curve for such data. One method is the Vincent technique, in which the total learning period for each

*S* is divided into equal trial parts (usually tenths), and the total number of correct responses in each part counted and then averaged for all *Ss*. Another method is the trials-to-criteria curve. In this method the number of trials required by each *S* to attain each successive criterion of 1, 2, 3 . . . *N* correct responses on a single trial is determined, and these values are averaged for all *Ss* for each criterion. The present method of describing the rote-learning process starts with a trials-to-criteria curve.

The trials-to-criteria curve is a series of points representing the earliest point in learning at which the average *S* achieved each criterion. If a representation is desired which combines maximum performance and earliest achievement of that maximum, no better method could be devised. But if it is assumed that learning involves a series of systematic fluctuations or oscillations, it can be seen that the trials-to-criteria method is "picking off" performance at the top of the oscillations. Melton (2) noted this a number of years ago, showing that, following attainment of each successive criterion, many *Ss* "fell back" to lower levels of performance. Now it would seem that if the series of points defining the trials-to-criteria curve represent the peak of fluctuations, a series of points showing the lowest or poorest performance after each peak might represent the nadir or bottom of the oscillations.

As illustrative data we have used the records for 100 *Ss* learning a difficult, ten-item serial list. This list was composed of nonsense syllables of low association value and of high intralist similarity. An average of 42.6 trials was required to learn the list to one perfect

<sup>1</sup> This work was done under Contract N7onr-45008, Project NR 154-057, between Northwestern University and the Office of Naval Research.

trial (3). The following data were obtained from the learning records.

1. Average number of trials required to reach each successive criterion.

2. Average lowest performance shown after attaining each successive criterion.

3. Average number correct on the trial just *before* the trial on which *S* attained each criterion. These values could be obtained for all criteria except one and two correct responses. Many *Ss* got two correct responses on the first anticipation trial, so that there is no "before" trial. Two *Ss* got three correct on the first anticipation trial, and one got four. Therefore, the *N* for the "before" trial for three correct responses is 98, for four correct 99, and for all others, 100.

4. Average number correct on the trial just *after* the trial on which *S* attained each successive criterion. All *Ss* are represented for these values up through six correct responses. After seven, *N* is 99, after eight, 98, and after nine, 86. These reduced *Ns* result from

*Ss*' "jumping" late criteria. For example, *S* may get eight correct on one trial and ten correct on the next, so that no record is available for the trial after attaining nine correct. Since these *Ss* who jump are usually fast-learning *Ss*, the mean number correct on the trial after achieving nine correct responses is probably slightly low. If *S* jumped from, say, two correct to four correct on successive trials, and if on the trial after getting four correct he again got four correct, he was recorded as having achieved four correct on the trial after reaching the criteria of two, three, and four correct.

These four sets of data are plotted on a single graph in Figure 1. The abscissa represents trials, and the ordinate the mean number correct. The upper dotted curve, therefore, simply represents a trials-to-criteria curve, and shows the usual smoothness. The open circles represent performance on the trial immediately after each criterion, and the closed circles performance on the trial just be-

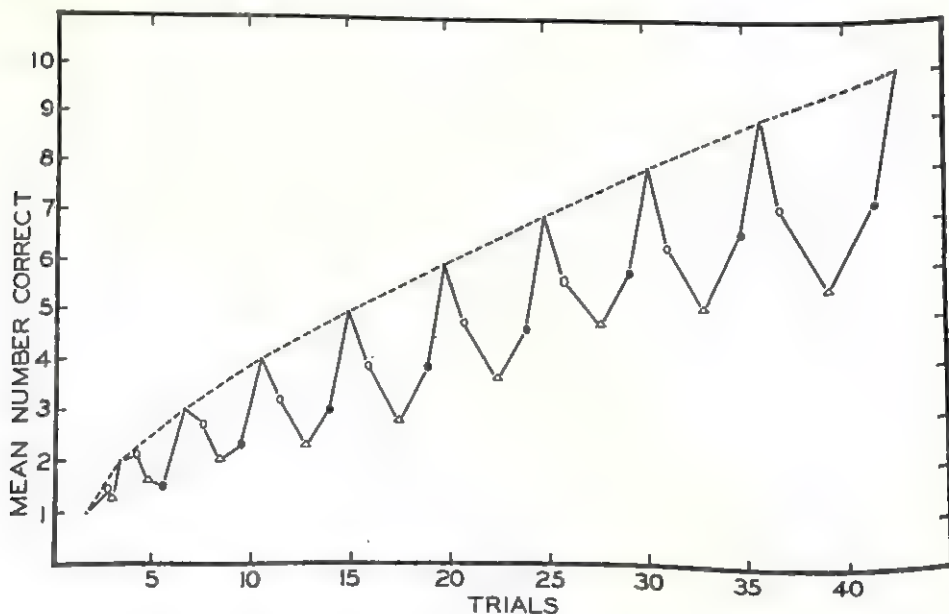


FIG. 1. Trials-to-criteria learning curve and a cyclical curve from the same data. See text for complete explanation.



fore each criterion is reached. For each criterion these are placed one trial after and one trial before, respectively, according to the abscissa scale. Finally, the points shown by the triangles represent the lowest performance shown after attaining each criterion; these triangles have been placed halfway between the before and after points.

It may be noted that the mean lowest performance after the criterion of two correct responses was achieved (the lowest point between two and three correct responses) is not as low as the "before" three correct performance. This apparent contradiction results again from S's jumping early criteria. Suppose S got 2, 5, and 6 items correct on three consecutive trials. His performance on the trial just before getting three correct is recorded as two. But, his lowest point after achieving two correct (the lowest point between two and three correct responses) is five correct. The "before" measure must always be lower than the criterion it precedes; the low point need not be.

The following characteristics of the solid-line curve may be noted:

1. Except after the criterion of one correct, performance falls at some point after attaining each criterion. As would be expected from plotting such curves for less difficult lists, the easier the list the longer are such falls delayed, i.e., they occur only after late criteria. If the average number of trials to learn is less than the number of items in the list, it is possible that no falls after any of the criteria would be observed.

2. The higher the criterion, the further the fall and the longer or more prolonged is this lower performance. If the amount of fall is changed to percentage of the immediately preceding criterion, however, the amount of fall so expressed is relatively constant for all criteria after the first two.

3. The performance on the trial just before attaining a criterion is consistently lower than on the trial just after reaching that criterion. However, the performances on the trial just after attaining a criterion and on the trial just before attaining the next are roughly equal.

It should be repeated that the list on which Figure 1 is based is a difficult one; easier lists will show less extreme fluctuations. Furthermore, with easier lists it will only be possible to get the low points between successive criteria and maintain relatively full Ns. These low points are critical for describing the bottom of the oscillations; the performance on trials just before and just after attaining each criterion can be omitted without changing the basic picture.

It seems evident that Figure 1 gives quite a different picture of the learning process from that of the typical learning curve (the dotted curve). However, it remains to be seen whether or not the cyclical curve has any useful or special theoretical significance. Those who use *oscillation* as a theoretical construct may find some comfort in these data; it is conceivable that certain physiological cycles might be shown to be related to the performance curve. An examination of the differences between fast and slow learners for the present data show that slow Ss fall further after reaching each criterion than do fast Ss. These slow Ss fall further on the trial immediately after each criterion and their low point is lower. However, their performance on the trial just before attaining each criterion is as high as the fast-learning Ss.

#### SUMMARY

It has long been known that group learning curves do not reflect the oscillations present in the learning records

of the individual S. A way of coordinating these oscillations for a group of Ss is presented for rote-learning data. The resulting graph shows systematic cycles in the performance curve.

## REFERENCES

1. HAYES, K. J. The backward curve: a method for the study of learning. *Psychol. Rev.*, 1953, 60, 269-275.
2. MELTON, A. W. The end-spurt in memorization curves as an artifact of the averaging of the individual curves. *Psychol. Monogr.*, 1936, 47, No. 2 (Whole No. 212). Pp. 119-134.
3. UNDERWOOD, B. J., & RICHARDSON, J. The influence of meaningfulness, intralist similarity, and serial position on retention. *J. exp. Psychol.*, 1956, 52, 119-126.

(Received August 27, 1956)

## ON PERCEPTUAL READINESS<sup>1</sup>

JEROME S. BRUNER

*Harvard University*

About ten years ago I was party to the publication of an innocent enough paper entitled "Value and Need as Organizing Factors in Perception." It was concerned with what at that time was the rather obscure problem of how extra-stimulus factors influenced perception, a subject then of interest to only a small band of us—Gardner Murphy, Nevitt Sanford, Muzafer Sherif, and a few others. Obviously, Professor Bor-ing is quite right about the mischievousness of the *Zeitgeist*, for the appearance of this paper seemed to coincide with all sorts of spirit-like rumblings within the world of psychology that were soon to erupt in a most unspirit-like torrent of research on this very topic—perhaps three hundred research reports and theoretical explications in the ten years since then. F. H. Allport (1) and M. D. Vernon (81) have each recently had a fresh look at the field, sorting out the findings and evaluating the theoretical positions, and they have done superb service. Their labors free me to pursue a more relaxed course. What I should like to do in this paper is to set forth what seem to me to be the outlines of an approach to perception congruent with this body of new (and often contradictory) findings and to sketch out what appear to me to be the persistent problems still outstanding.

<sup>1</sup> The present paper was prepared with the invaluable assistance of Mr. Michael Wallach. I also benefitted from the comments of Professors W. C. H. Prentice, Karl Pribram, and M. E. Bitterman, and from various associates at Princeton University, Kansas University, and the University of Michigan, where versions of this paper were presented.

### ON THE NATURE OF PERCEPTION

Perception involves an act of categorization. Put in terms of the antecedent and subsequent conditions from which we make our inferences, we stimulate an organism with some appropriate input and he responds by referring the input to some class of things or events. "That is an orange," he states, or he presses a lever that he has been "tuned" to press when the object that he "perceives" is an orange. On the basis of certain defining or criterial attributes in the input, what are usually called cues although they should be called clues (35), there is a selective placing of the input in one category of identity rather than another. The category need not be elaborate: "a sound," "a touch," "a pain," are also examples of categorized inputs. The use of cues in inferring the categorial identity of a perceived object, most recently treated by Bruner, Goodnow, and Austin (9) and by Binder (4), is as much a feature of perception as the sensory stuff from which percepts are made. What is interesting about the nature of the inference from cue to identity in perception is that it is in no sense different from other kinds of categorial inferences based on defining attributes. "That thing is round and nubbly in texture and orange in color and of such-and-such size—therefore an orange; let me now test its other properties to be sure." In terms of process, this course of events is no different from the more abstract task of looking at a number, determining that it is divisible only by itself and unity, and thereupon categorizing it in the class of prime numbers. So at the outset, it is evident



that one of the principal characteristics of perceiving is a characteristic of cognition generally. There is no reason to assume that the laws governing inferences of this kind are discontinuous as one moves from perceptual to more conceptual activities. In no sense need the process be conscious or deliberate. A theory of perception, we assert, needs a mechanism capable of inference and categorizing as much as one is needed in a theory of cognition.

Let it be plain that no claim is being made for the utter indistinguishability of perceptual and more conceptual inferences. In the first place, the former appear to be notably less docile or reversible than the latter. I may know that the Ames distorted room that looks so rectangular is indeed distorted, but unless conflicting cues are put into the situation, as in experiments to be discussed later, the room still looks rectangular. So too with such compelling illusions as the Miller-Lyer: in spite of knowledge to the contrary, the line with the extended arrowheads looks longer than the equal-length one with those inclined inward. But these differences, interesting in themselves, must not lead us to overlook the common feature of inference underlying so much of cognitive activity.

Is what we have said a denial of the classic doctrine of sense-data? Surely, one may argue (and Hebb [36] has done so effectively) that there must be certain forms of primitive organization within the perceptual field that make possible the differential use of cues in identity categorizing. Both logically and psychologically, the point is evident. Yet it seems to me foolish and unnecessary to assume that the sensory "stuff" on which higher order categorizations are based is, if you will, of a different sensory order than more evolved identities with which our perceptual world is normally peopled. To argue other-

wise is to be forced into the contradictions of Locke's distinction between primary and secondary qualities in perception. The rather bold assumption that we shall make at the outset is that all perceptual experience is necessarily the end product of a categorization process.

And this for two reasons. The first is that all perception is generic in the sense that whatever is perceived is placed in and achieves its "meaning" from a class of percepts with which it is grouped. To be sure, in each thing we encounter, there is an aspect of uniqueness, but the uniqueness inheres in deviation from the class to which an object is "assigned." Analytically, let it be noted, one may make a distinction, as Gestalt theorists have, between a pure stimulus process and the interaction of that stimulus process with an appropriate memory trace—the latter presumably resulting in a percept that has an identity. If indeed there is a "pure stimulus process," it is doubtful indeed that it is ever represented in perception bereft of identity characteristics. The phenomenon of a completely unplaceable object or event or "sensation"—even unplaceable with respect to modality—is sufficiently far from experience to be uncanny. Categorization of an object or event—placing it or giving it identity—can be likened to what in set theory is the placement of an element from a universe in a subset of that universe of items on the basis of such ordered dimensional pairs, triples, or  $n$ -tuples as man-woman, mesomorph-endomorph-ectomorph, or height to nearest inch. In short, when one specifies something more than that an element or object belongs to a universe, and that it belongs in a subset of the universe, one has categorized the element or object. The categorization can be as intersecting as "this is a quartz crystal goblet fashioned in Denmark," or as simple as "this is a

glassy thing." So long as an operation assigns an input to a subset, it is an act of categorization.

More serious, although it is "only a logical issue," is the question of how one could communicate or make public the presence of a nongeneric or completely unique perceptual experience. Neither language nor the tuning that one could give an organism to direct any other form of overt response could provide an account, save in generic or categorial terms. If perceptual experience is ever had raw, i.e., free of categorial identity, it is doomed to be a gem serene, locked in the silence of private experience.

Various writers, among them Gibson (26), Wallach (83), and Pratt (66), have proposed that we make a sharp distinction between the class of perceptual phenomena that have to do with the identity or object-meaning of things and the attributive or sensory world from which we derive our cues for inferring identities. Gibson, like Titchener (78) before him, urges a distinction between the visual field and the visual world, the former the world of attributive sense impressions, the latter of objects and things and events. Pratt urges that motivation and set and past experience may affect the things of the visual world but not the stuff of the visual field. And Wallach too reflects this ancient tradition of his Gestalt forebears by urging the distinction between a stimulus process pure and the stimulus process interacting with a memory trace of past experience with which it has made a neural contact on the basis of similarity. The former is the stuff of perception; the latter the finished percept. From shirtsleeves to shirtsleeves in three generations: we are back with the founding and founded content of the pre-Gestalt Gestalters. If one is to study the visual field freed of the things of the visual world, it becomes necessary—as Wallach implies—to free oneself of

the stimulus error: dealing with a percept not as an object or as a thing with identity, but as a magnitude or a brightness or a hue or a shape to be matched against a variable test patch.

If we have implied that categorizing is often a "silent" or unconscious process, that we do not experience a going-from-no-identity to an arrival-at-identity, but that the first hallmark of *any* perception is some form of identity, this does not free us of the responsibility of inquiring into the origin of categories. Certainly, Hebb (36) is correct in asserting like Immanuel Kant, that certain primitive unities or identities within perception must be innate or autochthonous and not learned. The primitive capacity to categorize "things" from "background" is very likely one such, and so too the capacity to distinguish events in one modality from those in others—although the phenomena of synesthesia would suggest that this is not so complete a juncture as it might seem; e.g., von Hornbostel (39). The sound of a buzz saw does rise and fall phenomenally as one switches room illumination on and off. The full repertory of innate categories—a favorite topic for philosophical debate in the 19th century—is a topic on which perhaps too much ink and too little empirical effort have been spilled. Motion, causation, intention, identity, equivalence, time, and space, it may be persuasively argued, are categories that must have some primitive counterpart in the neonate. And it may well be, as Piaget (65) implies, that certain primitive capacities to categorize in particular ways depend upon the existence of still more primitive ones. To identify something as having "caused" something else requires, first, the existence of an identity category such that the two things involved each may conserve identity in the process of "cause" producing "effect." Primitive or unlearned categories—a



matter of much concern to such students of instinctive behavior as Lashley (51) and Tinbergen (77)—remain to be explicated. In what follows, we shall rather cavalierly take them for granted. As to the development of more elaborated categories in terms of which objects are placed or identified, it involves the process of learning how to isolate, weigh, and use criterial attribute values, or cues for grouping objects in equivalence classes. It is only as mysterious, but no more so, than the learning of any differential discrimination, and we shall have occasion to revisit the problem later.

A second feature of perception, beyond its seemingly categorial and inferential nature, is that it can be described as varyingly veridical. This is what has classically been called the "representative function" of perception: what is perceived is somehow a representation of the external world—a metaphysical hodgepodge of a statement but one which we somehow manage to understand in spite of its confusion. We have long since given up simulacral theories of representation. What we generally mean when we speak of representation or veridicality is that perception is predictive in varying degrees. That is to say, the object that we *see* can also be *felt* and *smelled* and there will somehow be a match or a congruity between what we see, feel, and smell. Or, to paraphrase a younger Bertrand Russell, what we see will turn out to be the same thing should we take a "closer look" at it. Or, in still different terms, the categorial placement of the object leads to appropriate consequences in terms of later behavior directed toward the perceived object: it appears as an apple, and indeed it keeps the doctor away if consumed once a day.

Let it be said that philosophers, and notably the pragmatist C. S. Peirce, have been urging such a view for more

years than psychologists have taken their urgings seriously. The meaning of a proposition, as Peirce noted in his famous essay on the pragmatic theory of meaning (63), is the set of hypothetical statements one can make about attributes or consequences related to that proposition. "Let us ask what we mean by calling a thing *hard*. Evidently, that it will not be scratched by many other substances" (White, (84)). The meaning of a thing, thus, is the placement of an object in a network of hypothetical inference concerning its other observable properties, its effects, and so on.

All of this suggests, does it not, that veridicality is not so much a matter of representation as it is a matter of what I shall call "model building." In learning to perceive, we are learning the relations that exist between the properties of objects and events that we encounter, learning appropriate categories and category systems, *learning to predict and to check what goes with what*. A simple example illustrates the point. I present for tachistoscopic recognition two nonsense words, one a 0-order approximation to English constructed according to Shannon's rules, the other a 4-order approximation: YRULPZOC and VERNALIT. At 500 milliseconds of exposure, one perceives correctly and in their proper place about 48 per cent of the letters in 0-order words, and about 93 per cent of the letters in 4-order words. In terms of the amount of information transmitted by these letter arrays, i.e., correcting them for redundancy, the subject is actually receiving the same informational input. The difference in reportable perception is a function of the fact that the individual has learned the transitional probability model of what goes with what in English writing. We say that perception in one case is more "veridical" than in the other—the difference between 93 per



cent correct as contrasted with 48 per cent. What we mean is that the model of English with which the individual is working corresponds to the actual events that occur in English, and that if the stimulus input does not conform to the model, the resulting perception will be less veridical. Now let us drop the image of the model and adopt a more sensible terminology. Perceiving accurately under substandard conditions consists in being able to refer stimulus inputs to appropriate coding systems; where the information is fragmentary, one reads the missing properties of the stimulus input from the code to which part of the input has been referred. If the coding system applied does not match the input, what we read off from the coding system will lead to error and nonveridical perception. I would propose that perceptual learning consists not of making finer and finer discriminations as the Gibsons (27) would have us believe, but that it consists rather in the learning of appropriate modes of coding the environment in terms of its object character, connectedness, or redundancy, and then in allocating stimulus inputs to appropriate categorial coding systems.

The reader will properly ask, as Prentice (67) has, whether the notion of perceptual representation set forth here is appropriate to anything other than situations where the nature of the percept is not "clear"—perceptual representation under peripheral viewing conditions, in tachistoscopes, under extreme fatigue. If I am given a very good look at an object, under full illumination and with all the viewing time necessary, and end by calling it an orange, is this a different process from one in which the same object is flashed for a millisecond or two on the periphery of my retina with poor illumination? In the first and quite rare case the cues permitting the identification of the object are super-

abundant and the inferential mechanism operates with high probability relationships between cues and identities. In the latter, it is less so. The difference is of degree. What I am trying to say is that under *any* conditions of perception, what is achieved by the perceiver is the categorization of an object or sensory event in terms of more or less abundant and reliable cues. Representation consists of knowing how to utilize cues with reference to a system of categories. It also depends upon the creation of a system of categories-in-representation that fit the nature of the world in which the person must live. In fine, adequate perceptual representation involves the learning of appropriate categories, the learning of cues useful in placing objects appropriately in such systems of categories, and the learning of what objects are likely to occur in the environment, a matter to which we will turn later.

We have neglected one important feature of perceptual representation in our discussion: representation in perception of the space-time-intensity conditions of the external world. Perceptual magnitudes correspond in some degree to the metrical properties of the physical world that we infer from the nature of our perception. That is to say, when one line *looks* longer than another, it is likely to *be* longer as measured by the ruler. There are constant errors and sampling errors in such sensory representation, but on the whole there is enough isomorphism between perceiving without aids (psychology) and perceiving with aids (physics) to make the matter perennially interesting.

Is this form of representation subject to the kinds of considerations we have been passing in review? Does it depend upon categorizing activities and upon the construction of an adequate system of categories against which stimulus inputs can be matched? There is prob-

ably one condition where perceptual acts are relatively free of such influences, and that is in the task of discriminating simultaneously presented stimuli as alike or different—provided we do not count the “tuning of the organism” that leads one to base his judgment on one rather than another feature of the two stimuli. Ask the person to deal with one stimulus at a time, to array it in terms of some magnitude scale, and immediately one is back in the familiar territory of inferential categorizing. Prentice, in his able defense of formalism in the study of perception (67), seems to assume that there is a special status attached to perceptual research that limits the set of the observer to simple binary decisions of “like” and “different” or “present” and “absent,” and to research that also provides the subject with optimal stimulus conditions, and Graham (31) has recently expressed the credo that no perceptual laws will be proper or pure laws unless we reduce perceptual experimentation to the kinds of operations used in the method of constant stimuli.

There was at one time a justification for such a claim on the grounds that such is the best strategy for getting at the sensory-physiological processes that underlie perception. As we shall see in a later section, current work in neurophysiology brings this contention into serious doubt. In any case, the point must be made that many of the most interesting phenomena in sensory perception are precisely those that have been uncovered by departing from the rigid purism of the method of constants. I have in mind such pioneering studies as those of Stevens on sensory scales, where the organism is treated as an instrument whose sensory categorizations and scalar orderings are the specific object of study (74). Add to this the advances made by Helson on adaptation level (37) and by Volkman on the anchoring of sensory scales (82)—both

using the “sloppy” method of single stimuli—and one realizes that the nature of representation in perception of magnitudes is very much subject to categorizing processes, and to perceptual readiness as this is affected by subjective estimates of the likelihood of occurrence of sensory events of different magnitudes. Indeed, Helson’s law of adaptation level states that the subjective magnitude of a singly presented stimulus depends upon the weighted geometric mean of the series of stimuli that the subject has worked with, and the ingenious experiments of Donald Brown (7) have indicated that this adaptation level is influenced only by those stimuli that the subject considers to be within the category of objects being considered. Ask the subject to move a weight from one side of the table to the other with the excuse that it is cluttering up the table, and the weight does not serve as an anchor to the series, although it will show a discernible effect if it is directly included in the series being judged. In short, the category systems that are utilized in arraying magnitudes are also affected by the requirement of matching one’s model of the world to the actual events that are occurring—even if the categories be no more complicated than “heavy,” “medium,” and “light.”

The recent work of Stevens (75) on “the direct estimation of sensory magnitudes” highlights the manner in which veridicality in sensory judgment depends upon the prior learning of an adequate category set in terms of which sensory input may be ordered. Subjects are presented a standard tone of 1000 cps at 80 db. sound-pressure-level and are told that the value of this loudness is 10. Nine variable loudnesses all of the 1000 cps are then presented, varying 70 db. on either side of the standard, each one at a time being paired with the standard. “If the standard is called 10, what would you call



the variable? Use whatever numbers seem to you appropriate—fractions, decimals, or whole numbers." If one then compares the categorial judgments made with the sound pressure level of the various tones presented, using a log-log plot (log of the magnitude estimation against log of sound-pressure-level), the resulting function is a straight line, described by the empirical formula

$$L = kI^{0.3},$$

where  $L$  is loudness and  $I$  intensity. In short, categorial sorting of sensory magnitudes provides one with a mapping or representation of physical intensity. There are, to be sure, many problems connected with such a procedure, but the point remains: the magnitude categories in terms of which we scale sensory events represent a good fit to the physical characteristics of the world. Call this "veridicality" if you wish—although I do not see what is gained thereby; yet whatever one calls it, one must not lose sight of the fact that the judgments made are predictive of other features of the sensory inputs. Given the empirical conversion formula, one can predict from categorial judgment to physical meter readings.

To summarize, we have proposed that perception is a process of categorization in which organisms move inferentially from cues to categorial identity and that in many cases, as Helmholtz long ago suggested, the process is a silent one. If you will, the inference is often an "unconscious" one. Moreover, the results of such categorizations are representational in nature: they represent with varying degrees of predictive veridicality the nature of the physical world in which the organism operates. By predictive veridicality I mean simply that perceptual categorization of an object or event permits one to "go beyond" the properties of the object or event perceived to a prediction of other properties

of the object not yet tested. The more adequate the category systems constructed for coding environmental events in this way, the greater the predictive veridicality that results.

Doubtless, the reader will think of any number of examples of perceptual phenomena not covered by the simple picture we have drawn. Yet a great many of the classic phenomena are covered—psychophysical judgment, constancy, perceptual identification, perceptual learning, and so on. This will become clearer in the following sections. What must now be dealt with are the phenomena having to do with selectivity: attention, set, and the like.

#### CUE UTILIZATION AND CATEGORY ACCESSIBILITY

A fruitful way of thinking of the nature of perceptual readiness is in terms of the accessibility of categories for use in coding or identifying environmental events. Accessibility is a heuristic concept, and it may be defined in terms of a set of measures. Conceive of a person who is perceptually ready to encounter a certain object, an apple let us say. *How* he happens to be in this state we shall consider later. We measure the accessibility of the category "apples" by the amount of stimulus input of a certain pattern necessary to evoke the perceptual response "there is an apple," or some other standardized response. We can state the "minimum" input required for such categorization by having our observer operate with two response categories, "yes" and "no," with the likelihood of occurrence of apples and non-apples at 50:50, or by using any other definition of "maximum readiness" that one wishes to employ. The greater the accessibility of a category, (a) the less the input necessary for categorization to occur in terms of this category, (b) the wider the range of input characteristics that will be "accepted" as fitting



the category in question, (c) the more likely that categories that provide a better or equally good fit for the input will be masked. To put it in more ordinary language: apples will be more easily and swiftly recognized, a wider range of things will be identified or misidentified as apples, and in consequence the correct or best fitting identity of these other inputs will be masked. This is what is intended by accessibility.

Obviously, categories are not isolated. One has a category "apples," to be sure, but it is imbedded by past learning in a network of categories: "An apple a day keeps the doctor away" is one such category system. So too, are "apples are fruits" and other placements of an object in a general classification scheme. Predictive systems are of the same order: e.g., "The apple will rot if not refrigerated." We have spoken of these systems before as the "meaning" of an object. We mention them again here to indicate that though we speak analytically of separate or isolated categories as being accessible to inputs, it is quite obvious that category systems vary in accessibility as a whole.

It follows from what has just been said that the most appropriate pattern of readiness at any given moment would be that one which would lead on the average to the most "veridical" guess about the nature of the world around one at the moment—best guess here being construed, of course, as a response in the absence of the necessary stimulus input. And it follows from this that the most ready perceiver would then have the best chances of estimating situations most adequately and planning accordingly. It is in this general sense that the ready perceiver who can proceed with fairly minimal inputs is also in a position to use his cognitive readiness not only for perceiving what is before him but in foreseeing what is likely to

be before him. We shall return to this point shortly.

We must turn now to the question of cue utilization, the "strategies" in terms of which inferences are made (by the nervous system, of course) from cue to category and thence to other cues. I prefer to use the term strategy for several reasons. Perceiving, since it involves inference, rests upon a decision process, as Brunswik (17), Tanner and Swets (76) and others have pointed out. Even in the simplest threshold-measurement test, the subject has the task of deciding whether what he is seeing or hearing is noise only or signal-plus-noise. Given a set of cues, however presented, my nervous system must "decide" whether the thing is an airplane or a sea gull, a red or a green, or what not.

There appears, moreover, to be a sequence of such decisions involved in categorizing an object or event. A common-sense example will make this clear. I look across to the mantelpiece opposite my desk and see a rectangular object lying on it. If I continue this pursuit, subsequent decisions are to be made: is it the block of plastic I purchased for some apparatus or is it a book? In the dim light it can be either. I remember that the plastic is downstairs in one of the experimental rooms: the object "is" a book now, and I search for further cues on its dark red surface. I see what I think is some gold: it is a McGraw-Hill book, probably G. A. Miller's *Language and Communication* that I had been using late this afternoon. If you will, the process is a "bracketing" one, a gradual narrowing of the category placement of the object.

Let us attempt to analyze the various stages in such a decision sequence.

*a. Primitive categorization.* Before any more elaborate inferential activity can occur, there must be a first, "silent"

process that results in the perceptual isolation of an object or an event with certain characteristic qualities. Whether this is an innate process or one depending upon the prior construction of a cell-assembly, in the manner of Hebb (36), need not concern us. What is required simply is that an environmental event has been perceptually isolated and that the event is marked by certain spatio-temporal-qualitative characteristics. The event may have no more "meaning" than that it is an "object," a "sound," or a "movement."

*b. Cue search.* In highly practiced cases or in cases of high cue-category probability linkage, a second process of more precise placement based on additional cues may be equally silent or "unconscious." An object is seen with phenomenal immediacy as a "book" or an "ash tray." In such instances there is usually a good fit between the specifications of a category and the nature of the cues impinging on the organism—although "fit" and "probability of linkage" may stand in a vicarious relation to each other. Where the fit to accessible categories is not precise, or when the linkage between cue and category is low in probability in the past experience of the organism, the conscious experience of cue searching occurs. "What is that thing?" Here, one is scanning the environment for data in order to find cues that permit a more precise placement of the object. Under these circumstances, the organism is "open" to maximum stimulation, in a manner described below.

*c. Confirmation check.* When a tentative categorization has occurred, following cue search, cue search changes. The "openness" to stimulation decreases sharply in the sense that now, a tentative placement of identity having occurred, the search is narrowed for additional confirmatory cues to check this

placement. It is this feature of perceptual identification that Woodworth (85) in his paper on the "Reinforcement of Perception" speaks of as "trial-and-check." We shall speak of a selective gating process coming into operation in this stage, having the effect of reducing the effective input of stimulation not relevant to the confirmatory process.

*d. Confirmation completion.* The last stage in the process of perceptual identification is a completion, marked by termination of cue searching. It is characteristic of this state that openness to additional cues is drastically reduced, and incongruent cues are either normalized or "gated out." Experiments on the perception of incongruity (14), error (69), and the like (15), suggest that once an object has been categorized in a high-probability, good-fit category, the threshold for recognizing cues contrary to this categorization increases by almost an order of magnitude.

The question of fit between cue and category specification brings us to the key problem of the nature of categories. By a category we mean a rule for classifying objects as equivalent. The rule specifies the following about the instances that are to be comprised in the category.

*a.* The properties or *critical attributes* required of an instance to be coded in a given class.

*b.* The manner in which such attribute values are to be combined in making an inference from properties to category membership: whether conjunctively (e.g.,  $a_i$  and  $b_i$ ), relationally (e.g.,  $a_i$  bears a certain relation to  $b_i$ ), or disjunctively (e.g.,  $a_i$  or  $b_i$ ).

*c.* The weight assigned various properties in making an inference from properties to category membership.

*d.* The acceptance limits within which properties must fall to be critical. That



is to say, from what range of attribute values may  $a_i, b_i \dots k_i$  be drawn.

When we speak of rules, again it should be made clear that "conscious rules" are not intended. These are the rules that govern the operation of a categorizing mechanism.

The likelihood that a sensory input will be categorized in terms of a given category is not only a matter of fit between sensory input and category specifications. It depends also on the accessibility of a category. To put the matter in an oversimplified way, given a sensory input with equally good fit to two nonoverlapping categories, the more accessible of the two categories would "capture" the input. It is in this sense that mention was earlier made about the vicarious relationship between fit and accessibility.

We have already noted that the accessibility of categories reflects the learned probabilities of occurrence of events in the person's world. The more frequently in a given context instances of a given category occur, the greater the accessibility of the category. Operationally, this means that less stimulus input will be required for the instance or event to be categorized in terms of a frequently used category. In general, the type of probability we are referring to is not absolute probability of occurrence, where each event that occurs is independent of each other. Such independence is rare in the environment. Rather, the principal form of probability learning affecting category accessibility is the learning of contingent or transitional probabilities—the redundant structure of the environment. That either the absolute or the contingent probability of events makes a crucial difference in determining ease of perceptual identification is readily supported by research findings: in the former case by studies like those of Howes (40) and Solomon

and Postman (72), and in the latter by the work of Miller, Heise, and Lichten (62) and Miller, Bruner, and Postman (61).

But the organism to operate adequately must not only be ready for likely events in the environment, the better to represent them, and in order to perceive them quickly and without undue cognitive strain: it must also be able to search out unlikely objects and events essential to its maintenance and the pursuit of its enterprises. If I am walking the streets of a strange city and find myself hungry, I must be able to look for restaurants regardless of their likelihood of occurrence in the environment where I now find myself. In short, the accessibility of categories I employ for identifying the objects of the world around me must not only reflect the environmental probabilities of objects that fit these categories, but also reflect the search requirements imposed by my needs, my ongoing activities, my defenses, etc. And for effective search behavior to occur, the pattern of perceptual readiness during search must be realistic: tempered by what one is likely to find in one's perceptual world at that time and at that place as well as by what one seeks to find.

Let me summarize our considerations about the general properties of perception with a few propositions. The first is that *perception is a decision process*. Whatever the nature of the task set, the perceiver or his nervous system decides that a thing perceived is one thing and not another. A line is longer or shorter than a standard, a particular object is a snake and not a fallen branch, the incomplete word L\*VE in the context MEN L\*VE WOMEN is the word LOVE and not LIVE.

The second proposition is that *the decision process involves the utilization of discriminatory cues*, as do all decision processes. That is to say, the properties



of stimulus inputs make it possible to sort these inputs into categories of best fit.

Thirdly, *the cue utilization process involves the operation of inference.* Going from cue to an inference of identity is probably the most ubiquitous and primitive cognitive activity. The utilization of inference presupposes the learning of environmental probabilities and invariances relating cues to cues, and cues to behavioral consequences. Cue utilization involves various stages: a primitive step of isolating an object or event from the flux of environmental stimulation, stages of cue searching where the task is to find cues that can be fitted to available category specifications, a tentative categorization with more search for confirming cues, and final categorization, when cue searching is severely reduced.

Fourth, *a category may be regarded as a set of specifications* regarding what events will be grouped as equivalent—rules respecting the nature of criterial cues required, the manner of their combining, their inferential weight, and the acceptance limits of their variability.

Fifth, *categories vary in terms of their accessibility*, the readiness with which a stimulus input with given properties will be coded or identified in terms of a category. The relative accessibility of categories and systems of categories seems to depend upon two factors: the expectancies of the person with regard to the likelihood of events to be encountered in the environment; and the search requirements imposed on the organism by his needs and his ongoing enterprises. To use the functionalist's language, perceptual readiness or accessibility serves two functions: *to minimize the surprise value of the environment* by matching category accessibility to the probabilities of events in the world about one, and *to maximize the*

*attainment of sought-after objects and events.*

*Veridical perception*, so our sixth proposition would run, *consists of the coding of stimulus inputs in appropriate categories* such that one may go from cue to categorial identification, and thence to the correct inference or prediction of other properties of the object so categorized. Thus, veridical perception requires the learning of categories and category systems appropriate to the events and objects with which the person has commerce in the physical world. When we speak of the representative function of perception, we speak of the adequacy of the categorizing system of the individual in permitting him to infer the nature of events and to go beyond them to the correct prediction of other events.

Seventh, *under less than optimal conditions, perception will be veridical in the degree to which the accessibility of categorizing systems reflects the likelihood of occurrence of the events that the person will encounter.* Where accessibility of categories reflects environmental probabilities, the organism is in the position of requiring less stimulus input, less redundancy of cues for the appropriate categorization of objects. In like vein, nonveridical perception will be systematic rather than random in its error insofar as it reflects the inappropriate readiness of the perceiver. The more inappropriate the readiness, the greater the input or redundancy of cues required for appropriate categorization to occur—where "appropriate" means that an input is coded in the category that yields more adequate subsequent predictions.

#### MECHANISMS MEDIATING PERCEPTUAL READINESS

Having considered some of the most general characteristics of perceiving, particularly as these relate to the phenomena of perceptual readiness, we must

turn next to a consideration of the kinds of mechanisms that mediate such phenomena. Four general types of mechanisms will be proposed: *grouping and integration, access ordering, match-mismatch signaling, and gating*. They will be described in such a form that they may be considered as prototypes of neural mechanisms and, where possible, neurophysiological counterparts will be described briefly. Six years ago, Edward Tolman (79) proposed that the time was perhaps ripe for reconsidering the neural substrate of perception. Perhaps he was right, or perhaps even now the enterprise is somewhat premature. Yet, the body of perceptual data available makes it worth while to consider the kinds of mechanisms that will be required to deal with them. To use Hebb's engaging metaphor, it is worth while to build a bridge between neurophysiology and psychology provided we are anchored at both ends, even if the middle of the bridge is very shaky.

### *Grouping and Integration*

It is with the neural basis of the categorizing process that Hebb's *Organization of Behavior* (36) is principally concerned. Little is served by recapitulating his proposals here, for the reader will be familiar with the concise account in Chapters 4 and 5 of that book, where the concepts of cell assembly and phase sequence are set forth with a clarity that permits one to distinguish what is neurophysiological fact and what speculation. In essence, Hebb's account attempts to provide an anatomical-physiological theory of how it is that we distinguish classes of events in the environment, and how we come to recognize new events as exemplars of the once established classes. The theory seeks also to provide a mechanism for integration of sorting activity over time: the formation of phase sequences for the conservation of superordinate classes

of events and superordinate sequences. Basically, it is an associational or an "enrichment" theory of perception at the neural level, requiring that established neural associations facilitate perception of events that have gone together before. The expectancies, the centrally induced facilitations that occur prior to the sensory process for which they are appropriate, are learned expectancies based on the existence of frequency integrators. These frequency integrators may be neuroanatomical in the form of synaptic knobs, or they may be any process that has the effect of making activity in one locus of the brain increase or decrease the likelihood of activity in another. To be sure, Hebb's theory depends upon some broad assumptions about convergence of firing from area 17 outward, about synchronization of impulses, and about the manner in which reverberatory circuits can carry organization until the much slower process of anatomical change can take place. But this is minor in comparison with the stimulation provided by facing squarely the question of how the known facts of categorization and superordination in perception *could* be represented in the light of present knowledge.

While it is difficult indeed to propose a plausible neural mediator to account for category formation and the development of elaborated categorial systems (e.g., our knowledge of the relations between classes of events in the physical world which we manipulate in everyday life), it is less difficult to specify what such mechanisms must account for in perceptual behavior.

At the level of the individual category or cell assembly, the phenomena of object identity must be accounted for. Moreover, identity conservation or object constancy requires explanation in terms common with the explanation of identity. Experiments by Piaget (65)



suggest that the capacity to maintain the phenomenal identity of an object undergoing change is the hard-won result of maturation-and-learning. In connection with the later discussion of gating processes, we shall have occasion to consider the manner in which, at different stages in cue utilization, the required fit between an input and a cell assembly changes.

Where integration is concerned, there must be a process capable of conserving a record of the likely transitions and contingencies of the environment. The moment-to-moment programming of perceptual readiness depends upon such integrations. In short, the relation between classes of events is conserved in such a way as to be subject to change by learning. Several things can be guessed about integration processes. It is unlikely that it is a simple autocorrelation device. Clearly, the conceptions of transitional probabilities that are established in dealing with sequences of events show biases that no self-respecting autocorrelation computer would be likely to operate with. One of these is a strong and early tendency to treat events as nonindependent of each other over time. In the absence of evidence, or even in the presence of contrary evidence, humans—as their behavior has been observed in choice tasks, e.g., Estes (23), Goodnow (29)—treat random sequences of events as though they were governed by dependent probabilities. The spate of research on two-choice decision behavior has made us quite sharply aware of this characteristic of cognitive functioning. The typical pattern is the gambler's fallacy or, more properly, the negative recency effect. Given two equiprobable events whose occurrences are random, the repetition of one event progressively leads to the expectancy of the other. As in the elegant experiments of Jarvik (44) and Goodnow (29), the probability that a person will

predict one of two events increases directly as a function of the number of repetitions of the other event. Such behavior persists over thousands of opportunities for testing, and it appears under a variety of testing conditions (9).

The second feature of sequential probability integration mechanisms is that, in establishing a conception of the probability with which events will occur, the typical human subject will bias his estimate in terms of desired or feared outcomes. As in the experiments of Marks (60) on children and of Irwin (41) on adults, the subjectively estimated probability of strongly desired events will be higher per previous encountered occurrence than the estimated likelihood of less desired events. Quite clearly, then, the establishment of estimates depends upon more than frequency integrations biased by assumptions of nonindependence. The "something more" is a motivational or personality process, and we shall have more to say about it in considering phenomena of so-called "perceptual sensitization" and "perceptual defense."

#### *Access Ordering*

The term "accessibility" has been used in preceding pages to denote the ease or speed with which a given stimulus input is coded in terms of a given category under varying conditions of instruction, past learning, motivation, etc. It has been suggested, moreover, that two general sets of conditions affect accessibility: subjective probability estimates of the likelihood of a given event, and certain kinds of search sets induced by needs and by a variety of other factors.

Let us consider a few relevant facts about perception. The first of these is that the threshold of recognition for stimuli presented by visual, auditory, or other means is not only a function of



the time, intensity, or "fittingness" of the stimulus input, but also varies massively as a function of the number of alternatives for which the perceiver is set. The size of the expected array, to say it another way, increases the identification threshold for any item in the array. Typical examples of this general finding are contained in papers by Miller, Heise, and Lichten (62) and by Bruner, Miller, and Zimmerman (10). The actual shape of the function need not concern us, save that it is quite clear that it is not what one would expect from a simple binary system with a fixed channel capacity. What we are saying holds, of course, only for the case where the perceiver has learned that all the items in the expected array are (a) equiprobable and (b) independent, one of the other, in order of appearance.

The first hunch we may propose, then, about access-ordering mechanisms is that degree of accessibility of coding categories to stimulus inputs is related to regulation of the number of preactivated cell assemblies that are operative at the time of input. In an earlier paper (8), discussing factors that strengthen an hypothesis in the sense of making it more easily confirmable, I proposed that one of the major determinants of such strength was monopoly: where one and only one hypothesis is operative with no competing alternatives, it tends to be more readily confirmable. It is the same general point that is being made here. Accessibility, then, must have something to do with the resolution of competing alternatives.

As between two arrays of expected alternatives, each of the same size, we may distinguish between them in terms of the bias that exists in terms of expected likelihood of occurrence of each alternative. If one could characterize the expected alternatives in terms of probability values, one could conceive of the array ranging in values from a figure approaching 1.0 at

one extreme, to another approaching 0.0 at the other. The findings with respect to perceptual readiness for the alternatives represented in such an array are well known. For a constant-sized array, the greater the estimated likelihood of occurrence of an alternative, the more readily will the alternative be perceived or identified. This is known to be true for large arrays, such as the ensemble of known words in the English language, whose likelihood may be roughly judged by their frequency of occurrence in printed English (e.g., 40). It is not altogether clear that it is the case for arrays of expected alternatives that are within the so-called span of attention—i.e., less than seven or eight alternatives. That the principle holds for middling arrays of about 20 items has been shown by Solomon and Postman (72).

What is particularly interesting about change of accessibility, under conditions where estimates of the likelihood of occurrence of alternatives become biased, is that the biasing can be produced either by a gradual learning process akin to probability learning or by instruction. Thus, Bitterman and Kniffin (5), investigating recognition thresholds for taboo and neutral words, show that as the experiment progresses, there is a gradual lowering of threshold for the taboo words as the subject comes to expect their occurrence. Bruner and Postman (14) have similarly shown that repeated presentation of stimulus materials containing very low-probability incongruities leads to a marked decrease in threshold time required for recognizing the incongruous features. At the same time, both Cowen and Beier (20) and Postman and Crutchfield (70) have shown that if a subject is forewarned that taboo words are going to be presented, his threshold for them will tend to be lower than for neutral words, whereas it will be higher if no instruction is given. In short, preactivation of

cell assemblies—assuming for a moment that *degree of preactivation* is the mechanism that represents subjective estimates of likelihood of occurrence of an event—such preactivation can be produced by gradual learning or quantally by instruction. Moreover, biasing may be produced by the nature of the situation in which the perceiver is operating. A recent study by Bruner and Minturn (11) illustrates the point. Subjects are presented at brief exposure a broken capital B with a small separation between the vertical and the curved component of the letter so that it may be perceived as a B or as a 13. The manner in which it is reported is determined by whether the subject has previously been presented with letters or with numbers to recognize. In short, expectancy of one or the other context preactivates a related array of categories or cell-assemblies, not just a single, isolated one.

What the neural correlates of access ordering will look like is anybody's guess. Lashley (52) has remarked that, for all our searching, we have not located a specific memory trace—either in the form of a reverberatory circuit, a definite change in fiber size as proposed by J. Z. Young (88) and Eccles (21), a synaptic knob—in the manner of Lorente de No (57) or in any known form. To be sure, Penfield (64) has activated memories by punctate electrical stimulation of the cortex, but this is a long remove from a definition of the neural properties of the trace. For the time being, one does better to deal in terms of the formal properties that a trace system must exhibit than to rest one's psychological model on any neurophysiological or anatomical conception of the memory trace.

And, quite clearly, one of the formal properties of a trace system is that its elements vary in accessibility to stimulus input with the kinds of conditions we have considered. It is instructive to

note that when a theory of traces lacks this feature, it ceases to be useful in dealing with the wide range of perceptual categorizing phenomena of which we now have knowledge. Gestalt theory is a case in point. According to Köhler's view (48), a stimulus process "finds" its appropriate memory trace, resulting in identification of the stimulus process, on the basis of distinctive similarity between stimulus process and memory trace. The theory has been criticized, justly I think, for failing to specify the nature of this similarity save by saying that it is a neural isomorph of phenomenal similarity. But since similarity may be highly selective—two objects may be alike in color but differ in dozens of other respects—there is obviously some *tertium quid* that determines the basis of similarity. More serious still is the inability of such a theory to deal with the increased likelihood of categorization in terms of particular traces as a function of changes in search set or subjective likelihood estimates. The Bruner-Minturn results would require that, as between two traces with which a stimulus process may make contact, each equally "similar" to the stimulus, the stimulus process will make contact with the one having a higher probability of being matched by environmental events. This is interesting, but it is far from the spirit of Gestalt theory.

#### *Match-Mismatch Processes*

One may readily conceive of and, indeed, build an apparatus that will accept or reject inputs on the basis of whether or not they fulfill certain specifications. Selfridge (71) has constructed a machine to read letters, Fry (24) has one that will discriminate various phonemes, and Uttley (80) has constructed one that, like Tinbergen's graylag geese, will recognize the flying silhouette of a predator hawk. All such machines have



in common that they require a match between a stimulus input and various specifications required by the sorting mechanism of the machine.

In the examples just given, there is no consequence generated by whether a given input fulfills the specifications required by the identifying machine. It fits or it doesn't fit. But now let us build in two other features. The first is that the machine emit a signal to indicate how closely any given input comes to fulfilling the specifications required: either by indicating how many attributes the object has in common with the specifications, or by indicating how far off the mark on any given attribute dimension a given input is. The second is that the machine do something on the basis of these signals: to increase sensitivity if an object is within a given distance of specifications for a closer look, or to decrease it if the object is further than a certain amount from specifications, or to stop registering further if the input fits.

In short, one can imagine a nervous system that emits all-or-none match-mismatch signals or graded match-mismatch signals, and one can also imagine that these signals could then feed into an effector system to regulate activity relevant to continuing search behavior for a fitting object, or to regulate other forms of activity. MacKay (59) has recently proposed such a model.

We must return for a moment to an earlier discussion. In the discussion of cue utilization, a distinction was made between three phases of "openness" in cue search. The first was one in which a given input was being scanned for its properties so as to place it in one of a relatively large set of possible alternative categories. Here one would register on as many features of an object as possible. In a second stage, the input has been tentatively placed, and the search is limited to confirming or in-

firming criterial cues. Finally, with more definite placement, cue search is suspended and deviations from specification may even be "normalized." It is for the regulation of such patterns of search or cue utilization that some mechanism such as match-mismatch signaling is postulated.

Let it be said that while match-mismatch signaling-effector systems are readily conceivable and readily constructed, there is no knowledge available as to how a system like the nervous system might effect such a process. That there is feedback all over the system is quite apparent from its detailed anatomy, and this is the process out of which a larger-scale system such as we have described would be constructed.

### *Gating Processes*

The picture thus far presented is of a conceptual nervous system with a massive afferent intake that manages somehow to sort inputs into appropriate assemblies of varying accessibility. It seems unlikely that this is the nature of the nervous system, that there should be no gating or monitoring of stimulus input short of what occurs at higher centers. It is with this more peripheral form of screening of inputs that we shall now be concerned.

It has long been known that the concept of the "adequate stimulus" could not simply be defined as a change in environmental energy sufficient to stimulate a receptor. For quite evidently, a stimulus could be peripherally adequate in this sense and not be "centrally" adequate at all, either in eliciting electrical activity in the cortex or in producing a verbal report of a change in experience by the subject. Indeed, the very nature of such complex receptor surfaces as the retina argues against such a simple notion of "adequate stimulus." For the reactivity of even a retinal cell at the fovea seems to be "gated" by the state



of stimulation of neighboring cells. Thus, if cells A, B, and C lie next each other in that order in a row, stimulation of B suppresses the sensitivity of C. If A now be stimulated, B is suppressed and C is released or heightened in sensitivity. So even at the level of the first synapse of a sensory system, there is mediation *outward* or gating from internuncial to receptor cells that programs the nature of the input that can come into the sensory system. And to be sure, there are many phenomena in perception itself that speak for this same kind of gating. When we are fixated upon the vase in the Rubin reversible figure, the background recedes, is less surfacy, and in general seems to provide a generally less centrally adequate form of sensory input. So too with the studies of Yokoyama (87) and Chapman (19) where subjects, set to report on one of several attributes of briefly presented stimuli, accomplished their selective task with a loss of ability to discriminate on the attributes for which they had not been set. We shall propose that such phenomena are very likely mediated by a gating process which "filters" input before ever it reaches the cortex.

There is now a growing body of neurophysiological evidence that part of this screening process is relegated to peripheral levels of the nervous system—even as far out as the second synapse of specialized sensory systems. In an earlier paper I used the rather fanciful phrase that "perception acts sometimes as a welcoming committee and sometimes as a screening committee." It now appears that both these committees are closer to the entrance port than previously conceived.

Consider first the evidence of Kuffler and Hunt (50) on so simple a "reflex" as the stretch reflex of the biceps femoris muscle of the cat in an isolated spinal nerve-muscle preparation. Recall a little anatomy first. Muscle tissue con-

tains special cells called spindles that are receptors in function, discharging with contraction or stretch of the muscle in which they are imbedded. The muscle itself is innervated by an efferent nerve trunk emerging from the ventral horn of the spinal cord and, in turn, an afferent nerve travels to the dorsal root of the spinal cord. According to the classical law of Bell and Magendie, the ventral root of the spinal cord carries efferent-motor impulses down to the muscles, while the dorsal root carries sensory impulses up to the cord. Now, it has been known for a long time that the presumed efferent nerve going to muscles carries fibers of large and of small diameter. A quarter-century ago Eccles and Sherrington showed that the ventral nerve branch supplying the biceps femoris of the cat shows a "striking division of the fibers into two diameter groups" (49), one group centering around  $5\mu$  in diameter, the other around  $15$  or  $16\mu$ . The large fibers are, of course, fast conductors, the small ones slow. Leksell (55) has shown that stimulation of the slow-conducting smaller fibers did not cause detectable contractions or propagated muscle impulses. When the larger and fast-conducting fibers are stimulated, the usual motor-unit twitch occurred. Kuffler and Hunt (50) state that, in the lumbosacral outflow, about  $\frac{2}{3}$  of the fibers are of the large-diameter, fast-conduction type; the other third are of the small type that in mammalia are "ineffective in directly setting up significant muscular contraction." There has been much speculation about what these fibers are there for, and the answer is now fairly clear. It is revolutionary in its implications and brings deeply into question both the classical Bell-Magendie law and the simplistic notion of the reflex arc on which so much of American learning theory is based.

It is this. The small fibers of the

presumably motor trunk go to the spindle cells and the activity in these fibers serve to modulate or gate the receptivity of these specialized sensory endings. For example, if the small-diameter fibers are firing into the muscle spindle it may speed up the amount of firing from this cell into the afferent nerve that is produced by a given amount of stretch tension on the muscle. We need not go into detail here. It suffices to note that the state of presumed motor discharge does not simply innervate the muscle; it also regulates the amount and kind of kinesthetic sensory discharge that the sensory cells in the muscle will send back to the central nervous system. Instead of thinking of a stimulus-response reflex arc, it becomes necessary even at this peripheral level to think of the efferent portion of the arc acting back on sensory receptors to change the nature of the stimulus that can get through.

Two additional pieces of evidence on gating mechanisms at higher levels of integration may be cited. Where vision is concerned, Granit (32) has recently shown that pupillary changes produced by the ciliary muscle of the eye create changes in the pattern of firing of the retina: changes in muscular state working its way back through the nervous system into the visual system and back outward to the retina. There is also evidence of gating working from the visual system backward in the opposite direction: during binocular rivalry, the nondominant eye shows a less sensitive pupillary reflex than the dominant eye.

Finally, we may cite the recent evidence of Hernandez-Péon, Scherrer, and Jouvett (38) working in Magoun's laboratory, work confirmed by analogous findings of Golambos, Sheatz, and Vernier (28) at the Walter Reed Hospital. If one stimulates the cat with auditory clicks, it is possible to record an evoked spike potential from the cochlear nucleus. Repetition of the clicks leads

to a gradual diminution of the evoked potential, as if the organism were adapting. It is quite extraordinary that such adaptation should be registered as far out peripherally as the cochlear nucleus, which is, after all, only the second synapse of the VIIIth nerve. Now, if the clicks are previously used as conditioned stimuli signaling shock, the diminution of the evoked potential no longer occurs upon repetition of the clicks. Evidence that the response from the brain is not being produced by the muscular activity produced by the click as a conditioned stimulus is provided by the fact that the same kind of effects are obtained from cats with temporarily induced muscular paralysis. Further, if one take a cat whose cochlear nucleus is still firing upon click stimulation and introduce a mouse into its visual field, the clicks no longer evoke a spike potential. A fish odor or a shock to the paw has the same effect of inhibiting spike potentials at the cochlear nucleus, if these distracting stimuli occur concurrently with the click. "Distraction" or "shifting of attention" appears to work its way outward to the cochlear nucleus.<sup>2</sup>

Perhaps the foregoing account has been needlessly detailed on the side of neurophysiology. Yet, the interesting implications of the findings for perceptual theory make such an excursion worth while. That the nervous system accomplishes something like gating is quite clear, even without the neurophysiological evidence. The data of behavior are full of examples, and the phenomena of attention require some such mechanism to be explained. Indeed, it is quite

<sup>2</sup> Since the above was written, evidence has been presented by Golambos indicating that efferently controlled inhibition operates as far out to the periphery as the hair cells of the organ of Corti and fibers carrying such inhibitory impulses have been traced as far centrally as the superior olivary nucleus—not very far, but a start.



clear that the nervous system must be capable of more selective gating than physiology has yet been able to discover. That is to say, there must be a filter somewhere in the cat's nervous system that will "pass" the squeak of the mouse in the Hernandez-Péon experiment but not the cough of the experimenter. And it is to this problem that we turn now.

I would propose that one of the mechanisms operative in regulating search behavior is some sort of gating or filtering system. In the preceding section, it was proposed that the "openness" of the first stage of cue utilization, the "selectivity" of the second stage, and the "closedness" of the third stage were probably regulated by a match-mismatch mechanism. What may be proposed here is that the degree of "openness" or "closedness" to sensory input during different phases of cue utilization is likely effected by the kind of gating processes we have been considering. How these work in intimate detail is far from known, yet the work of the last years in neurophysiology suggests that we are drawing closer to an answer.

Having considered some general properties of perception and some possible mechanisms underlying these, we turn now to some selected problems in perception better to explore the implications of what has thus far been proposed.

#### ON FAILURE OF READINESS

From the foregoing discussion, it is clear that veridical perception under viewing or listening conditions that are less than ideal depends upon a state of perceptual readiness that matches the probability of occurrence of events in the world of the perceiver. This is true, of course, only in a statistical sense. What is most likely to occur is not necessarily what will occur, and the perceiver whose readiness is well matched to the likelihoods of his environment

may be duped. In Farquhar's handsome seventeenth-century phrase: "I cou'd be mighty foolish, and fancy myself mighty witty; reason still keeps its Throne—but it nods a little, that's all." The only assurance against the nodding of reason or probability, under the circumstances, is the maintenance of a flexibility of readiness: an ability to permit one's hypotheses about what it is that is to be perceptually encountered to be easily infirmed by sensory input. But this is a topic for later.

There appear to be two antidotes to nonveridical perception, two ways of overcoming inappropriate perceptual readinesses. The one is a re-education of the misperceiver's expectancies concerning the events he is to encounter. The other is the "constant close look." If the re-education succeeds in producing a better match between internal expectancies and external event-probabilities, the danger of misperception under hurried or substandard conditions of perceiving is lessened. But the matter of re-educating perceptual expectancies is complex. For where consequences are grave, expectancy concerning what may be encountered does not change easily, even with continued opportunity to test the environment. In this concluding section we shall consider some of the factors that contribute to states of perceptual "unreadiness" that either fail to match the likelihood of environmental events or fail to reflect the requirements of adjustment or both.

Before turning to this task, a word is in order about the "constant close look" as an antidote to inappropriate perceptual readiness. There is for every category of objects that has been established in the organism a stimulus input of sufficient duration and cue redundancy such that, if the stimulus input fits the specifications of the category, it will eventually be correctly perceived as an exemplar of that category. With enough



time and enough testing of defining cues, such "best fit" perceiving can be accomplished for most but not all classes of environmental events with which the person has contact. There are some objects whose cues to identity are sufficiently equivocal so that no such resolution can be achieved, and these are mostly in the sphere of so-called interpersonal perception: perceiving the states of other people, their characteristics, intentions, etc., on the basis of external signs. And since this is the domain where misperception can have the most chronic if not the most acute consequences, it is doubtful whether a therapeutic regimen of "close looking" will aid the misperceiver much in dealing with more complex cue patterns. But the greatest difficulty rests in the fact that the cost of close looks is generally too high under the conditions of speed, risk, and limited capacity imposed upon organisms by their environment or their constitutions. The ability to use minimal cues quickly in categorizing the events of the environment is what gives the organism its lead time in adjusting to events. Pause and close inspection inevitably cut down on this precious interval for adjustment.

### *Inappropriate Categories*

Perhaps the most primitive form of perceptual unreadiness for dealing with a particular environment is the case in which the perceiver has a set of categories that are inappropriate for adequate prediction of his environment. A frequently cited example of such a case is Bartlett's account (3) of the African visitors in London who perceived the London bobbies as especially friendly because they frequently raised their right hand, palm forward, to the approaching traffic. The cue-category inference was, of course, incorrect, and they should have identified the cue as a signal for stopping traffic. The example, however,

is not particularly interesting because it is a transient phenomenon, soon corrected by instruction.

A more interesting example, because it is far less tractable, is provided by second-language learning and the learning of a new phonemic system. Why is it, we may ask, that a person can learn the structure of a new language, its form classes, morphemes, lexemes, and so on, but still retain a "foreign accent" which he cannot, after a while, distinguish from the speech flow of native speakers around him? And why is it that a person learning a new language can follow the speech of a person with his own kind of foreign accent more readily than he can follow a native speaker? The answer lies, I think, in the phenomenon of postcategorization sensory gating: once an utterance has been "understood" or decoded in appropriate categories, on the basis of some of the diacritica of the speech flow, the remaining features are assimilated or normalized or screened out. The phonemic categories that are used, moreover, are modifications of those in the first language of the speaker. Normalization is in the direction of these first-language phonemic categories. It is only by a special effort that, after having achieved adequate comprehension of the second language, one can remain sensorially "open" enough to register on the deviation between his own phonemic pattern and that of native speakers. And since there is common categorization of the "meaning" of utterances by the native speaker and the fluent foreigner, there is no built-in incentive for the foreigner to maintain a cognitively strainful regimen of attending further to speech sounds.

Lenneberg (56) has recently shown the difficulties involved in learning new modes of categorizing such continua as chromatic colors. He taught subjects various nonsense languages, explaining

to them that the words were Hopi names for colors and that their task was to learn what colors they stood for. His stimulus materials were graded Munsell colors going in a circle from *brown*, through *green*, through *blue*, through *pink*, and then back to *brown*. A standardizing group was used to find the frequency distribution of color naming over the circle when the English color names mentioned above were used. Experimental groups, six in number, were then run, each being exposed to the use of the nonsense color names "as these are used by the Hopi." Then they were tested on their usage of the names. A first group was taught the nonsense words with exact correspondence to the usage found for the standardizing group on *brown*, *blue*, *green*, and *pink*. The other groups were given distorted usage training—distorted from English usage. The distortions were both in the slopes of the frequency of usage and in the points on the color continua where the highest usage frequencies fell. That is to say, the mode of a distribution in some cases would fall at a color which in English had no specific name, or fall between two English categories.

The principal results of the experiment are these. If the reference and probability relationship is the same for a nonsense language as it is for English, relearning is very rapid. The slightest deviation from this correspondence increases difficulty of learning quite markedly. It is disturbing either to shift the center of the categories on the color continuum or to change the shape of the frequency-of-calling functions, even when these are made *more* determinative (i.e., rectilinear) than they normally are. A shift in the shape of the frequency-of-calling functions is more disruptive than a shift in placement on the color continuum. What is quite striking is that a highly determinative frequency-of-calling function can be

learned much more rapidly than one in which there is a gradual transition in color naming from one color to another on the color continuum.

Now, I suspect that the difficulty in learning a set of neighboring categories with a state of equivocality prevailing in the area between the "typical instances" of each category comes precisely from the tendency to normalize in the direction of the center of one category or the other. If there is a sharp transition between one color category and another, this tendency aids learning. If the transition is gradual, it hinders it. For it is noteworthy, as in the experiment of Bruner, Postman, and Rodrigues (16) that equivocal colors are readily subject to assimilation in the direction of expected value.

It is perhaps in the realm of social perception, where the problem of validating one's categorizations is severe, that one finds the most striking effects of inappropriate category systems. What is meant here by validation is the testing of the predictions inherent in a categorization. If, on the basis of a few cues of personal appearance, for example, one categorizes another person as dishonest, it is extremely difficult in most cases to check for the other cues that one would predict might be associated with instances of this category. There is either a delay or an absence of opportunity for additional cue checking. Moreover, there is also the likelihood, since cues themselves are so equivocal in such a case, that available equivocal signs will be distorted in such a manner as to confirm the first impression. It is much as in the experiments of Asch (2) and of Haire and Grunes (33) on the formation of first impressions, where later cues encountered are cognitively transformed so as to support the first impression. The reticence of the man we categorize as dishonest is seen as "caginess," the "honest" man's reticence



is seen as "integrity" and "good judgment."

It is perhaps because of this difficulty of infirming such categorial judgments that an inappropriate category system can be so hard to change. The slum boy who rises to the top in science can change his categories for coding the events of the physical world quite readily. He has much more difficulty in altering the socially related category system with which he codes the phenomena of the social world around him.

### *Inappropriate Accessibility Ordering*

Perhaps the most noticeable "perceptual unreadiness" comes from interference with good probability learning by wishes and fears. I have in mind the kind of distorted expectancies that arise when the desirability or undesirability of events distorts the learning of their probability of occurrence. The experiments of Marks (60) and of Irwin (41), cited earlier, are simplified examples of the way in which desired outcomes increase estimates of their likelihood of occurrence. Certain more persistent general personality tendencies also operate in this sphere. It is indeed the case that some people are readier to expect and therefore quicker to perceive the least desirable event among an array of expected events, and others the most desired. This is quite clearly a learned adjustment to the events one is likely to encounter, even if it may be supported by temperamental characteristics. How such learning occurs, and why it is so resistant to correction by exposure to environmental events, are hardly clear. But one matter that becomes increasingly clear is that before we can know much about how appropriate and inappropriate perceptual readiness is produced, we shall have to know much more about how organisms learn the probabilistic structure of their environments. This is a point that Brunswik

has made for some years (17), and it is one that is now being taken seriously by such students of probability learning as Bush and Mosteller (18), Bruner, Goodnow, and Austin (9), Estes (23), Galanter and Gerstenhaber (25), Hake and Hyman (34), Edwards (22), and others.

There is another important feature of learning that affects perceptual readiness. It has to do with the range of alternatives for which organisms learn to be set perceptually. Put the matter this way. It is a matter of common observation that some people are characteristically tuned for a narrow range of alternatives in the situations in which they find themselves. If the environment is banal in the sense of containing only high probability events and sequences or, more properly, events and sequences that are strongly expected, then the individual will do well and perceive with a minimum of pause for close looking. But should the environment contain unexpected events, unusual sequences, then the result will be a marked slowdown in identification and categorizing. Cue search must begin again. We speak of such people as "rigid" or "stuck." George Klein's work (46) on shifting category judgments suggests that, in general, people who are not able to shift categorization under gradually changing conditions of stimulation tend also to show what he describes as "over-control" on other cognitive and motivational tasks. At the other extreme is specialization upon diversity, and how such specialization is learned is equally puzzling. I can perhaps best illustrate the phenomenon by a commonly observed pattern found in subjects in tachistoscopic experiments. There are subjects who show rather high thresholds of identification generally, and who seem to be "weighing" the stimulus in terms of a wide array of interpretive categories. Jenkin (45) has recently described such perception as "rational-



ized," the subject describing what he sees as "like a so-and-so" rather than, as in the "projective" response, reporting it "as a so-and-so." It is as if the former type of response involved a greater cue searching of stimulus inputs for a fit to a wide range of things that it "could be." It is also very likely that premature sensory gating occurs in individuals with a tendency to be set for a minimum array of alternatives, leading them into error. The topic is one that bears closer investigation. To anyone who has had much experience in observing subjects in tachistoscopic work, it seems intuitively evident that there are large and individual differences possibly worth examining here.

We come finally to the vexing problem of "perceptual defense"—the manner in which organisms utilize their perceptual readiness to ward off events that are threatening but about which there is nothing they can do. There has been foolish and some bitter ink spilled over this topic, mostly because of a misunderstanding. The notion of perceptual defense does not require a little homuncular ego, sitting behind a Judas-eye, capable of ruling out any input that is potentially disruptive—as even so able a critic as F. H. Allport (1) seems to think. Any preset filtering device can do all that is required.

Let me begin with the general proposition that failure to perceive is most often not a *lack* of perceiving but a matter of *interference* with perceiving. Whence the interference? I would propose that the interference comes from categorizations in highly accessible categories that serve to block alternative categorizations in less accessible categories. As a highly speculative suggestion, the mechanism that seems most likely to mediate such interference is probably the broadening of category acceptance limits when a high state of readiness to perceive prevails; or, in the

language of the preceding section, the range of inputs that will produce a match signal for a category increases in such a way that more accessible categories are likely to "capture" poor-fitting sensory inputs. We have already considered some evidence for increase in acceptance limits under high readiness conditions: the tendency to see a red four of clubs as either a four of diamonds or a four of clubs, with color-suit relationship rectified (14), the difficulty of spotting reversed letters imbedded in the middle of a word (69), and so on.

Let us examine some experimental evidence on the role of interference in perceptual failure. Wyatt and Campbell (86) have shown that if a subject develops a wrong hypothesis about the nature of what is being presented to him for perception at suboptimal conditions, the perception of the object in terms of its conventional identity is slowed down. This observation has been repeated in other studies as well. Postman and Bruner (68), for example, have shown that if a subject is put under pressure by the experimenter and given to believe that he is operating below standard, then he will develop premature hypotheses that interfere with correct perception of the word stimuli being presented to him. The authors refer to "perceptual recklessness" as characterizing the stressed subjects in contrast to those who operated under normal experimental conditions. It may well be, just in passing, that stress has not only the specific effect of leading to premature, interfering hypotheses but that it disrupts the normal operation of match-mismatch signaling systems in the nervous system. Unpublished studies from our own laboratory carried out by Bruner, Postman, and John (15) have shown the manner in which subjects misperceive low-probability contingencies in terms of higher probability categories. For example, a subject in the experi-

mental group is shown tachistoscopically a picture of a discus thrower, wound up and ready to throw. In his balancing arm and placed across the front of him is a large bass viol. A control subject is shown the same picture, the exact space filled by the bass viol now being occupied by the crouching figure of a track official with his back to the camera. The brightness, shading, and area of the viol and the official are almost identical. Subjects begin by identifying the first flash of the picture as an athlete with a shadow across him. The subjects faced with the incongruous picture then go on with reasonable hypotheses—including the hypothesis of a crouching human figure, "probably an official," as one subject put it—and in the process of running through the gamut of likely hypotheses, correct perception is interfered with. It will not surprise you if I report that the threshold for the incongruous stimulus picture is markedly higher than that for the more conventional one.

Hypotheses and states of readiness may interfere with correct perception in yet another way: by creating a shifting "noise" background that masks the cues that might be used for identifying an environmental event. At the common-sense level this can best be illustrated by reference to perceptual-motor learning where kinesthetic cues are of importance. In teaching a person how to cast a fly, it is necessary for him to guide his forward delivery by feeling the gentle pressure release that occurs when the line reaches the end of its uncurving on the backcast. If your flycasting pupil is too eager to spot this cue, he will be rather tense, and his own muscular tension will mask the gentle pressure release that he must use as a signal.

A good instance is provided by the experiment of Goodnow and Pettigrew (30) at Harvard. It is concerned with the ability of subjects to perceive a regu-

larity in a sequence of events—a very simple regularity, like the alternation left-right-left-right. . . . The experiment is done on a conventional two-armed bandit, the subject having the task of betting on whether a light will appear on the left or on the right. The task is simple. A subject is first given some pretraining, in one of four pretraining groups. One is given pretraining by learning a simple alternation pattern of payoff, another is trained to find the payoff all on one side (not easy for all subjects), a third is trained to find the pattern LLRLLR . . . , and a final group is given no pretraining. Following the pretraining and without pause, all subjects are given a series of 60 choices in which the payoff is randomly arranged, the two sides totaling out to 50:50. Immediately following this random phase, and again without pause, the payoffs now go into a stage of simple alternation, LRLR. . . . How long does it take the subject to perceive the regularity of the final temporal pattern? The speed of discovery depends, it turns out, upon the kinds of behavioral hypotheses a subject develops during the phase of random payoff. If he develops any regularity of response—like win-stay-lose-shift or win-shift-lose-stay—then he will quickly spot the new pattern. Pretraining on a constant one-side payoff or on single alternation both produce such regularity, and both forms of pretraining produce equally good results—the subject requiring but eight or nine exposures to the pattern introduced after the random phase to begin responding without error. No pretraining, or pretraining on the pattern LLRLLR . . . , does not produce the regularity of response required. Instead, the subject works on odd and constantly shifting hypotheses during the random period. When the single-alternation regularity is introduced, the result is a marked reduction in ability to spot the new pat-



tern—some subjects failing to discover the pattern in 200 trials. What we are dealing with here is interference—hypotheses and responses serve to mask the regularity of events in the environment. In order for an environmental regularity to be perceived, there has to be a certain amount of steadiness in the hypotheses being employed and in the response pattern that is controlled by it. Short of this, masking and clumsy perceptual performance results.

Now what has all this to do with "perceptual defense"? The concept was introduced some years ago by Postman and myself as a description of the phenomenon of failure to perceive and/or report material known by independent test to be regarded as inimical by the subject. It was proposed (13) that there was a hierarchy of thresholds, and that an incoming stimulus could be responded to without its reaching the level of reportable experience—as in the McGinnies (58) and Lazarus and McCleary (54) studies, where autonomic response followed presentation of a potentially traumatic stimulus without the subject's being able to give a report of the nature of the stimulus. The study of Bricker and Chapanis (6) threw further light on the concept of a hierarchy of thresholds by demonstrating that, though subjects could not report spontaneously on the identity of the shock syllables used by Lazarus and McCleary, they could guess them well in excess of chance if given a restricted choice regarding what word had been presented. I would like to propose two additional factors that might lead to a failure of perception of emotionally negative material.

It is conceivable that the estimates of probability of occurrence of disvalued events are, in some individuals, reduced—essentially the obverse of what was observed in the experiments of Marks (60) and Irwin (41), where probability

estimates were inflated by desirability. If accessibility is decreased by such disvaluation, then a cognitive counterpart of what is clinically called "repression" can be posited. It is known, however, that not everyone shows this tendency to be unready for objects and events that are anxiety-arousing. Others seem to *inflate* their estimate of the likelihood of occurrence of inimical events. Certainly one finds clinical evidence for such a pattern among anxiety neurotics. In an early paper, Postman and Bruner (68) described two types of performance with respect to known anxiety-producing stimuli, defense and vigilance, the former a heightened threshold of identification for such stimuli, the latter a lowered threshold. In a carefully designed experiment contrasting the performance of clinically diagnosed "intellectualizers" and "repressors," Lazarus, Eriksen, and Fonda (53) have shown that the former group indeed are faster in recognizing negatively charged material than they are in recognizing neutral material, while the latter show the reverse tendency. Again, I find it necessary to revert to a point made earlier. I do not think that we are going to get much further ahead in understanding hyper- and hyporeadiness for encountering anxiety-evoking stimuli short of doing studies of the learning of environmental probabilities for sequences containing noxious and beneficial events.

One additional mechanism that may be operative in lowering or generally in altering readiness to perceive material that in some way may be threatening. I hesitate to speak of it in detail, since it is of such a speculative order, and do so only because some experiments suggest themselves. It is this. Conceivably, categories for classes of objects that are pain-arousing are set up with narrow acceptance limits for stimulus inputs related to them. That is to say, what we speak of as "repression" may be the



establishment of very narrow category limits that prevent the evocation of match signals for inputs that do not fit category specifications very precisely. I am mindful that as far as autonomic reactivity is concerned potentially traumatic stimuli work in quite the reverse direction. If anything, a wide range of objects, appropriate and inappropriate, arouse autonomic reactions, without leading to verbalizable report concerning the categorial identity of the eliciting objects. Yet it is conceivable that with respect to one kind of threshold (autonomic) the acceptance limits are broad, and with respect to another (reportable awareness) very narrow. I think it would be worth while in any case to investigate the acceptance limits of inimical stimulus inputs by altering the characteristics of objects so that, in essence, one gets a generalization gradient for recognition. My guess is that the gradient will be much steeper for anxiety-arousing stimuli than for neutral ones. All that remains is to do the experiment.

Finally, it may also be the case that category accessibility reflects the instrumental relevance of the environmental events they represent. There is evidence that the recognition threshold for noxious objects about which one can do something is lower than normal, whereas for ones about which nothing instrumental can be done, the threshold is higher. That is to say, words that signal a shock that can be avoided show lowered thresholds, words signaling unavoidable shock show a threshold rise. One may well speculate whether the instrumental relevance of objects is not a controlling factor in guiding the kind of search behavior that affects category accessibility. The problem needs much more thorough investigation than it has received.

We have touched on various conditions that might lead a person to be inappropriately set for the events he must perceive easily and quickly in his en-

vironment. Many other studies could be mentioned. But the intention has not been to review the rather sprawling literature in the field, but to propose some possible mechanism affecting readiness so that research might be given a clearer theoretical direction.

### CONCLUSIONS

We have been concerned in these pages with a general view of perception that depends upon the construction of a set of organized categories in terms of which stimulus inputs may be sorted, given identity, and given more elaborated, connotative meaning. Veridical perception, it has been urged, depends upon the construction of such category systems, categories built upon the inference of identity from cues or signs. Identity, in fine, represents the range of inferences about properties, uses, and consequences that can be predicted from the presence of certain criterial cues.

Perceptual readiness refers to the relative accessibility of categories to afferent stimulus inputs. The more accessible a category, the less the stimulus input required for it to be sorted in terms of the category, given a degree of match between the characteristics of the input and the specifications of the category. In rough form, there appear to be two general determinants of category accessibility. One of them is the likelihood of occurrence of events learned by the person in the course of dealing with the world of objects and events and the redundant sequences in which these are imbedded. If you will, the person builds a model of the likelihood of events, a form of probability learning only now beginning to be understood. Again in rough terms, one can think of this activity as achieving a minimization of surprise for the organism. A second determinant of accessibility is the requirements of search dictated by need states

and the need to carry out habitual enterprises such as walking, reading, or whatever it is that makes up the round of daily, habitual life.

Failure to achieve a state of perceptual readiness that matches the probability of events in one's world can be dealt with in one of two ways: either by the relearning of categories and expectancies, or by constant close inspection of events and objects. Where the latter alternative must be used, an organism is put in the position of losing his lead time for adjusting quickly and smoothly to events under varying conditions of time pressure, risk, and limited capacity. Readiness in the sense that we are using it is not a luxury, but a necessity for smooth adjustment.

The processes involved in "sorting" sensory inputs to appropriate categories involve cue utilization, varying from sensorially "open" cue searching under relative uncertainty, to selective search for confirming cues under partial certainty, to sensory "gating" and distortion when an input has been categorized beyond a certain level of certainty.

Four kinds of mechanisms are proposed to deal with known phenomena of perceptual categorizing and differential perceptual readiness: *grouping and integration*, *access ordering*, *mismatch signal utilization*, and *gating*. The psychological evidence leading one to infer such processes were examined and possible neurological analogues considered. The processes are conceived of as mediators of categorizing and its forms of connectivity, the phenomena of differential threshold levels for various environmental events, the guidance of cue search behavior, and lastly, the phenomena of sensory inhibition and "filtering."

Finally, we have considered some of the ways in which failure of perceptual readiness comes about—first, through a failure to learn appropriate categories

for sorting the environment and for following its sequences, and second, through a process of interference whereby more accessible categories with wide acceptance limits serve to mask or prevent the use of less accessible categories for the coding of stimulus inputs. The concept of "perceptual defense" may be re-examined in the light of these notions.

In conclusion, it seems appropriate to say that the ten years of the so-called New Look in perception research seem to be coming to a close, with much empirical work accomplished—a great deal of it demonstrational, to be sure, but with a promise of a second ten years in which hypotheses will be more rigorously formulated and, conceivably, neural mechanisms postulated, if not discovered. The prospects are anything but discouraging.

#### REFERENCES

1. ALLPORT, F. H. *Theories of perception and the concept of structure*. New York: Wiley, 1955.
2. ASCH, S. E. *Social psychology*. New York: Prentice-Hall, 1952.
3. BARTLETT, F. C. *Remembering*. Cambridge, England: Cambridge Univer. Press, 1932.
4. BINDER, A. A statistical model for the process of visual recognition. *Psychol. Rev.*, 1955, **62**, 119-129.
5. BITTERMAN, M. E., & KNEFFIN, C. W. Manifest anxiety and "perceptual defense." *J. abnorm. soc. Psychol.*, 1953, **48**, 248-252.
6. BRICKER, P. D., & CHAPANIS, A. Do incorrectly perceived tachistoscopic stimuli convey some information? *Psychol. Rev.*, 1953, **60**, 181-188.
7. BROWN, D. R. Stimulus similarity and the anchoring of subjective scales. *Amer. J. Psychol.*, 1953, **66**, 199-214.
8. BRUNER, J. S. Personality dynamics and the process of perceiving. In R. R. Blake & G. V. Ramsey (Eds.), *Perception: an approach to personality*. New York: Ronald, 1951. Pp. 121-147.
9. BRUNER, J. S., GOODNOW, J. J., & AUSTIN, G. A. *A study of thinking*. New York: Wiley, 1956.



10. BRUNER, J. S., MILLER, G. A., & ZIMMERMAN, C. Discriminative skill and discriminative matching in perceptual recognition. *J. exp. Psychol.*, 1955, 49, 187-192.
11. BRUNER, J. S., & MINTURN, A. L. Perceptual identification and perceptual organization. *J. gen. Psychol.*, 1955, 53, 21-28.
12. BRUNER, J. S., & POSTMAN, L. Emotional selectivity in perception and reaction. *J. Pers.*, 1947, 16, 69-77.
13. BRUNER, J. S., & POSTMAN, L. Perception, cognition, and behavior. *J. Pers.*, 1949, 18, 14-31.
14. BRUNER, J. S., & POSTMAN, L. On the perception of incongruity: a paradigm. *J. Pers.*, 1949, 18, 206-223.
15. BRUNER, J. S., POSTMAN, L., & JOHN, W. Normalization of incongruity. Research memorandum. Cognition Project, Harvard Univer., 1949.
16. BRUNER, J. S., POSTMAN, L., & RODRIGUES, J. Expectation and the perception of color. *Amer. J. Psychol.*, 1951, 64, 216-227.
17. BRUNSWIK, E. *Systematic and representative design of psychological experiments*. Berkeley: Univer. of California Press, 1949.
18. BUSH, R. R., & MOSTELLER, C. F. *Stochastic models for learning*. New York: Wiley, 1955.
19. CHAPMAN, D. W. Relative effects of determinate and indeterminate Aufgaben. *Amer. J. Psychol.*, 1932, 44, 163-174.
20. COWEN, E. L., & BEIER, E. G. The influence of "threat expectancy" on perception. *J. Pers.*, 1951, 19, 85-94.
21. ECCLES, J. C. *The neurophysiological basis of mind*. Oxford: Oxford Univer. Press, 1953.
22. EDWARDS, W. The theory of decision making. *Psychol. Bull.*, 1954, 51, 380-417.
23. ESTES, W. K. Individual behavior in uncertain situations: an interpretation in terms of statistical association theory. In R. M. Thrall, C. H. Coombs, & R. L. Davis (Eds.), *Decision processes*. New York: Wiley, 1954. Pp. 127-137.
24. FRY, D. P., & DENES, P. Mechanical speech recognition. In W. Jackson (Ed.), *Communication theory*. New York: Academic Press, 1953.
25. GALANTER, E., & GERSTENHABER, M. On thought: extrinsic theory of insight. *Amer. Psychologist*, 1955, 10, 465.
26. GIBSON, J. J. *The perception of the visual world*. Boston: Houghton Mifflin, 1950.
27. GIBSON, J. J., & GIBSON, E. J. Perceptual learning: differentiation or enrichment? *Psychol. Rev.*, 1955, 62, 32-41.
28. GOLAMBOS, R., SHEATZ, G., & VERNIER, V. G. Electrophysiological correlates of a conditioned response in cats. *Science*, 1956, 123, 376-377.
29. GOODNOW, J. J. Determinants of choice-distribution in two-choice situations. *Amer. J. Psychol.*, 1955, 68, 106-116.
30. GOODNOW, J. J., & PETTIGREW, T. E. Some difficulties in learning a simple pattern of events. Paper presented at annual meeting of the East. Psychol. Ass., Atlantic City, March, 1956.
31. GRAHAM, C. H. Perception and behavior. Presidential address to the East. Psychol. Ass., Atlantic City, March, 1956.
32. GRANIT, R. *Receptors and sensory perception*. New Haven: Yale Univer. Press, 1955.
33. HAIRE, M., & GRUNES, W. F. Perceptual defenses: processes protecting an organized perception of another personality. *Hum. Relat.*, 1950, 3, 403-412.
34. HAKE, H. W., & HYMAN, R. Perception of the statistical structure of a random series of binary symbols. *J. exp. Psychol.*, 1953, 45, 64-74.
35. HARPER, R. S., & BORING, E. G. Cues. *Amer. J. Psychol.*, 1948, 61, 119-123.
36. HEBB, D. O. *The organization of behavior*. New York: Wiley, 1949.
37. HELSON, H. Adaptation-level as a basis for a quantitative theory of frames of reference. *Psychol. Rev.*, 1948, 55, 297-313.
38. HERNANDEZ-PÉON, R., SCHERRER, R. H., & JOUVET, M. Modification of electric activity in the cochlear nucleus during "attention" in unanesthetized cats. *Science*, 1956, 123, 331-332.
39. HORNOSTEL, E. M. VON. Unity of the senses. *Psyche*, 1926, 7, 83-89.
40. HOWES, D. On the interpretation of word frequency as a variable affecting speed of recognition. *J. exp. Psychol.*, 1954, 48, 106-112.
41. IRWIN, F. W. Stated expectations as functions of probability and desirability of outcomes. *J. Pers.*, 1953, 21, 329-335.
42. ITTLESON, W. H. *The Ames demonstrations in perception*. Princeton, N. J.: Princeton Univer. Press, 1952.
43. JARRETT, J. Strategies in risk-taking situations. Unpublished doctor's dissertation, Harvard Univer. Library, 1951.
44. JARVIK, M. E. Probability learning and a negative recency effect in the serial an-



- tipication of alternative symbols. *J. exp. Psychol.*, 1951, 41, 291-297.
45. JENKIN, N. Two types of perceptual experience. *J. clin. Psychol.*, 1956, 12, 44-49.
  46. KLEIN, G. S. The personal world through perception. In R. R. Blake & G. V. Ramsey (Eds.), *Perception: an approach to personality*. New York: Ronald, 1951. Pp. 328-355.
  47. KOHLER, I. Rehabilitation in perception. Published separately in three parts, in German, in *Die Pyramide*, 1953, Heft 5, 6, and 7 (Austria). Translated by Henry Gleitman and edited by J. J. Gibson. Privately circulated by the editor.
  48. KÖHLER, W. *Dynamics in psychology*. New York: Liveright, 1940.
  49. KUFFLER, S. W., HUNT, C. C., & QUILLIAN, J. P. Function of medullated small-nerve fibers in mammalian ventral roots: efferent muscle spindle innervation. *J. Neurophysiol.*, 1951, 14, 29-54.
  50. KUFFLER, S. W., & HUNT, C. C. The mammalian small nerve fibers: a system for efferent nervous regulation of muscle spindle discharge. *Proc. Assoc. Res. Nerv. Ment. Dis.*, 1952, Vol. 30.
  51. LASHLEY, K. S. Experimental analysis of instinctive behavior. *Psychol. Rev.*, 1938, 45, 445-471.
  52. LASHLEY, K. S. In search of the engram. *Symp. Soc. Exp. Biol.*, 1950, 4, 454-482.
  53. LAZARUS, R. S., ERIKSEN, C. W., & FONDA, C. P. Personality dynamics and auditory perceptual recognition. *J. Pers.*, 1951, 19, 471-482.
  54. LAZARUS, R. S., & MCCLEARY, R. A. Autonomic discrimination without awareness: a study of subception. *Psychol. Rev.*, 1951, 58, 113-222.
  55. LEKSELL, L. The action potential and excitatory effects of the small ventral root fibers to skeletal muscles. *Acta Physiol. Scand.*, 1945, 10, Suppl. 31.
  56. LENNEBERG, E. H. An empirical investigation into the relationship between language and cognition. Unpublished doctoral dissertation, Harvard Univer. Library, 1956.
  57. LORENTE DE NO, R. Transmission of impulses through cranial motor nuclei. *J. Neurophysiol.*, 1939, 2, 402-464.
  58. MCGINNIES, E. Emotionality and perceptual defense. *Psychol. Rev.*, 1949, 56, 244-251.
  59. MACKAY, D. M. Toward an information-flow model of human behavior. *Brit. J. Psychol.*, 1956, 47, 30-43.
  60. MARKS, R. W. The effect of probability, desirability, and "privilege" on the state of expectations of children. *J. Pers.*, 1951, 19, 332-351.
  61. MILLER, G. A., BRUNER, J. S., & POSTMAN, L. Familiarity of letter sequences and tachistoscopic identification. *J. gen. Psychol.*, 1954, 50, 129-139.
  62. MILLER, G. A., HEISE, G. A., & LICHTEN, W. The intelligibility of speech as a function of the context of the test materials. *J. exp. Psychol.*, 1951, 41, 329-335.
  63. PEIRCE, C. S. How to make our ideas clear. *Popular Sci. Mon.*, 1878, 12, 286-302.
  64. PENFIELD, W. Memory mechanisms. *Arch. Neurol. & Psychiat.*, 1952, 67, 178-191.
  65. PIAGET, J. *Play, dreams, and imitation in childhood*. New York: Norton, 1951.
  66. PRATT, C. C. The role of past experience in visual perception. *J. Psychol.*, 1950, 30, 85-107.
  67. PRENTICE, W. C. H. Paper read at the Symposium on Conceptual Trends in Psychology, at Amer. Psychol. Ass., New York, September, 1954.
  68. POSTMAN, L., & BRUNER, J. S. Perception under stress. *Psychol. Rev.*, 1948, 55, 314-323.
  69. POSTMAN, L., BRUNER, J. S., & WALK, R. D. The perception of error. *Brit. J. Psychol.*, 1951, 42, 1-10.
  70. POSTMAN, L., & CRUTCHFIELD, R. S. The interaction of need, set, and stimulus structure in a cognitive task. *Amer. J. Psychol.*, 1952, 65, 196-217.
  71. SELFRIDGE, O. Pattern recognition and learning. Memorandum of Lincoln Laboratory, Massachusetts Institute of Technology, 1955.
  72. SOLOMON, R. L., & POSTMAN, L. Frequency of usage as a determinant of recognition thresholds for words. *J. exp. Psychol.*, 1952, 43, 195-201.
  73. SMITH, J. W., & KLEIN, G. S. Cognitive control in serial behavior patterns. Dittied manuscript, available from author, 1951.
  74. STEVENS, S. S. Chapter I in S. S. Stevens (Ed.), *Handbook of experimental psychology*. New York: Wiley, 1951.
  75. STEVENS, S. S. The direct estimation of sensory magnitudes—loudness. *Amer. J. Psychol.*, 1956, 69, 1-25.

76. TANNER, W. P., JR., & SWETS, J. A. A decision-making theory of human detection. *Psychol. Rev.*, 1954, 61, 401-409.
77. TINBERGEN, N. *The study of instinct*. Oxford: Oxford Univer. Press, 1951.
78. TITCHENER, E. B. *A beginner's psychology*. New York: Macmillan, 1916.
79. TOLMAN, E. C. Discussion. *J. Pers.*, 1949, 18, 48-50.
80. TUTTLE, A. M. *The conditional probability of signals in the nervous system*. Radar Research Establ., British Ministry of Supply, Feb., 1955.
81. VERNON, M. D. *A further study of visual perception*. Cambridge, England: Cambridge Univer. Press, 1952.
82. VOLKMANN, J. In M. Sherif & J. H. Rohrer (Eds.), *Social psychology at the crossroads*. New York: Harpers, 1951.
83. WALLACH, H. Some considerations concerning the relation between perception and cognition. *J. Pers.*, 1949, 18, 6-13.
84. WHITE, M. *The age of analysis*. New York: New American Library, 1955.
85. WOODWORTH, R. S. Reenforcement of perception. *Amer. J. Psychol.*, 1947, 60, 119-124.
86. WYATT, D. F., & CAMPBELL, D. T. On the liability of stereotype or hypothesis. *J. abnorm. soc. Psychol.*, 1951, 46, 496-500.
87. YOKOYAMA, J. Reported in E. G. Boring, *A history of experimental psychology*. (2nd Ed.). New York: Appleton-Century, 1954.
88. YOUNG, J. Z. *Doubt and certainty in science*. Oxford: Oxford Univer. Press, 1951.

(Received June 4, 1956)

# THE PSYCHOLOGICAL REVIEW

## ON THE PSYCHOPHYSICAL LAW<sup>1</sup>

S. S. STEVENS

*Harvard University*

After dealing his critics their deserts, Fechner concluded his polemic of 1877 with a defiant five-line *Nachwort* (10).

The tower of Babel was never finished because the workers could not reach an understanding on how they should build it; my psychophysical edifice will stand because the workers will never agree on how to tear it down.

For nearly a century the critics have pecked at the edifice in the disorganized manner that inspired no fear in Fechner. They have contended that sensation is not measurable, and that it is; that Weber's law is not true, and that it is; that the just noticeable difference is not a proper unit, and that it is; that the logarithmic law is not found in experiment, and that it is. As James put it, Fechner's critics flailed about, "smiting his theories hip and thigh and leaving not a stick of them standing . . .," but here we are a hundred years later still discussing Fechner.

The lesson of history is that a bold and plausible theory that fills a scientific need is seldom broken by the impact of contrary facts and arguments. Only with an alternative theory can we hope to displace a defective one.

The purpose here is to try to do just

that—to try to show that there is a general psychophysical law relating subjective magnitude to stimulus magnitude, and that this law is simply that equal stimulus ratios produce equal subjective ratios. On numerous perceptual continua, direct assessments of subjective magnitude seem to bear an orderly relation to the magnitude of the stimulus. To a fair first-order approximation, the ratio scales constructed by "direct" methods (as opposed to the indirect procedures of Fechner) are related to the stimulus by a power function of one degree or another. Evidence for this fact is piling up under the impact of ratio-scaling procedures whose development over the past three decades has given a "new look" to psychophysics. The fruits of these new pursuits have implications and consequences for the traditional issues in the area of psychophysics, as well as for certain related issues concerned with category judgments, time-order errors, and scaling theory.

That the Fechnerian philosophy of indirect measurement will fade quietly away as soon as it has been shown what direct methods can achieve is scarcely to be counted on, for elaborations and applications of this philosophy pervade important segments of the psychological activity. Curiously enough, modern Fechnerians are found less frequently among the psychophysicists than among the psychometricians and scale construc-

<sup>1</sup> This research was supported partly by a grant from the National Science Foundation, partly by Contract Nonr-1866 (15) with the Office of Naval Research (Project Nr142-201, Report PNR-188). Reproduction for any purpose of the U. S. Government is permitted.



tors. It is they who sometimes assume that "equally often noticed differences are equal." As elaborated in Thurstone's "law of comparative judgment," this principle is essentially Fechner's principle: namely, that the unit of measurement is given by resolving power. If not explicitly, at least by implication, this philosophy of indirect measurement asserts that all we can know about magnitude is what confusion tells us. Variability becomes the measure of things, and the mean is meaningless. But on an important class of those simpler psychological continua where these notions are testable, we can show that equally often noticed differences are *not* equal, and that a scale proportional to psychological magnitude is not achieved by procedures that try to transform variabilities, discrimininal dispersions, or confusions into units of measure.

Before we consider this point further we will need to draw a distinction between two classes of continua.

## TWO CLASSES OF CONTINUA

Perceptual continua divide themselves into two general classes (64). The nature of this division is suggested in a general way by the traditional dichotomy between quantity and quality. Continua having to do with *how much* belong to what we have called Class I, or prothetic; continua having to do with *what kind* and *where* (position) belong to Class II, or metathetic. Class I seems to include, among other things, those continua on which discrimination is mediated by an additive or prothetic process at the physiological level (53, 67). An example is loudness, where we progress along the continuum by adding excitation to excitation. Class II includes continua on which discrimination is mediated by a physiological process that is substitutive, or metathetic. An example is pitch, where we progress along the continuum by substituting ex-

citation for excitation, i.e., by changing the locus of excitation.

We know too little about physiological mechanisms to say whether all Class I continua are based on additive processes or all Class II continua on substitutive processes, but in those instances where the facts seem clear the parallels between function and physiology are at least suggestive. Until our knowledge stands improved, it is perhaps best to classify the perceptual continua by the pragmatic criteria of the way they behave and to hope that any uniformities we can discover will lead to deeper insights into basic mechanisms.

Four functional criteria are relevant to the distinction between prothetic and metathetic continua, although they have not all been tested with equal thoroughness. These four criteria concern the subjective size of the j.n.d., the form of category rating-scales, the time-order error, and hysteresis. It is likely, of course, that other criteria will be discovered if we continue to look for them.

1. *Subjective size of the j.n.d.* I list this criterion first because it is the growth in the subjective size of the j.n.d. as we go up the scale on a Class I or prothetic continuum that seems, in a sense, to "explain" at least two of the other three criteria. Fechner proposed that scales can be constructed by counting off just noticeable differences. The implication is that the sensation produced by a stimulus 50 j.n.d.'s above threshold is half as great as that produced by a stimulus 100 j.n.d.'s above threshold. The hard fact of the matter is that if the typical subject were confronted with two such stimuli on a Class I continuum he would assert with certainty that the ratio between the two sensations is greater than two, because scales obtained by summing j.n.d.'s are nonlinearly related to ratio scales of subjective magnitude (33, 52). But more about this in a later section. The

point here is that on Class I continua the j.n.d.'s are not equal in subjective size.

On continua of Class II (metathetic) the j.n.d.'s turn out to be approximately equal in subjective size when measured by magnitude scales of the continuum. The linearity between the mel scale of subjective pitch and the j.n.d. scale for frequency is a case in point (56) and it seems reasonable to suppose that the same linear relation is approximated on such continua as position and inclination.

2. *Category rating-scales.* A category rating-scale is the function obtained when a subject judges a set of stimuli in terms of a set of categories labeled either by numbers or by adjectives. The form of these scales is different on the two kinds of continua. As shown by studies of a dozen perceptual dimensions (64), the category scales on continua of Class I are concave downward when plotted against a ratio scale of the subjective magnitude. Category scales on continua of Class II *may be* linear when so plotted.

The chief factor that produces non-linearity in the category scales of Class I is variation in the subject's sensitivity to differences. Near the lower end of the scale where discrimination is good the categories tend to be narrow, and by consequence the slope of the function is steep. Near the upper end, where a given stimulus difference is less easy to detect, the categories broaden and the slope declines. Only on Class II continua, where sensitivity (measured in subjective units) remains relatively constant, is it ordinarily possible to produce category scales that are linearly related to subjective magnitude.

Prothetic continua of Class I on which category scales have proved nonlinear include apparent length, area, numerosness, duration, heaviness, lightness, brightness and loudness. Metathetic

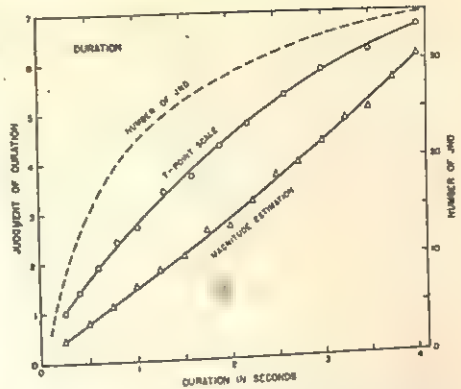


FIG. 1. J.n.d. scale, category scale, and ratio scale for apparent duration. *Triangles*: Mean judgments of 12 subjects who estimated the apparent duration of a white noise. Stimuli were presented in a different irregular order to each subject. *Circles*: Mean category judgments made by 16 subjects on a scale from 1 to 7. The end stimuli were presented at the outset to indicate the range, and each subject judged each duration twice in a random order. Stimulus spacing was adjusted to produce a "pure" category scale. *Dashed line*: Summated j.n.d.'s, right-hand ordinate.

continua of Class II on which category scales have proved more or less linear include visual position, inclination, proportion, and pitch. On both types of continua alterations in the form of the category scale may be produced by a variety of factors, including stimulus spacing (or relative frequency of presentation), landmarks, and differential familiarity.

By means of an iterative series of experiments we can achieve a stimulus spacing that will neutralize the effects of spacing and produce a "pure" category scale, uncontaminated by the subject's expectations regarding the frequency with which he should name the various categories. An example of an approximation to a pure category scale is shown in Fig. 1 (circles). This experiment was a follow-up to the experiments on duration reported earlier (64). Guided by the outcome of the earlier studies, I so spaced the stimuli that the 16 sub-

jects tended to name the various categories equally often. The proof that a pure category scale has been approximated lies in the fact that if we use the curve in Fig. 1 to determine what spacing of the stimuli will mark off equal distances along the ordinate, no significant change in the spacing is called for. (The technique of experimental iteration is a powerful device for neutralizing certain kinds of bias. For other examples of its use see [57, 64, 66].)

The iterated experiments on duration culminating in that shown in Fig. 1 converge on a category scale that is typical of those obtained on continua of Class I (prothetic). We note that it is concave downward, as expected, and that it contrasts sharply with the ratio scale (triangles) determined by asking

12 subjects to estimate directly the apparent duration of the stimulus. The upward curvature of the ratio scale for subjective duration agrees with the results of independent experiments in which the method of fractionation was used to determine the duration that appeared to be half as long as a standard duration (19, 50).

The curves in Fig. 1 tell us that on a prothetic continuum the typical subject is unable to equalize the intervals on his category scale, even when instructed to do so. He fails for the basic reason that his ability to tell one magnitude from another varies over the scale and affects the width of his categories. Since he can easily tell 0.5 sec. from 1.0 sec. he tends to put them in different categories; since he can only with difficulty tell 3.5 sec. from 4.0 sec. he tends to put them in the same category. Any procedure that tries directly or indirectly to get the subject to partition a prothetic continuum into equal intervals seems bound to fail in the general case.

For purposes of comparison, the function obtained by counting up j.n.d.'s is plotted in Fig. 1 as a dashed curve. This curve, based on Woodrow's (73) report that the relative j.n.d. for duration is of the order of 10 to 12.5 per cent, shows the form Fechner's law predicts for the scale of apparent duration.

**3. Time-order error.** The so-called "time-error" discovered by Fechner has been pursued in theory and experiment for many decades (23). It refers to the fact that the second of two equal stimuli tends to be judged greater than first. We have reason to believe that a systematic time-order error (as Guilford prefers to call it) is characteristic of judgments on Class I continua. On Class II continua we neither expect it nor do we generally find it.

This hypothesis was first suggested to me by the troubles we met when we

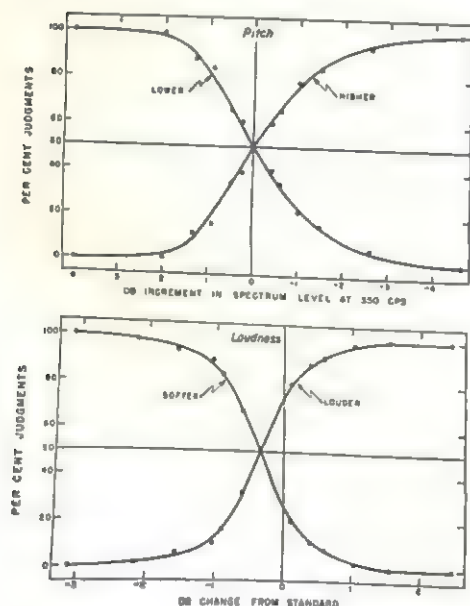


FIG. 2. Psychometric functions from 50 subjects obtained with recorded tests for ability to discriminate the pitch and loudness of bands of noise. Each test contained 110 test items consisting of a standard noise (2 seconds) followed immediately by a comparison noise (2 seconds). The band of noise used in the pitch test was shifted up or down the frequency scale. Note that there is a "time error" for loudness but none for pitch.



tried to supply the armed forces with recorded tests designed to assess pitch and loudness discrimination for bands of noise (29). Karlin and I discovered that it was easy enough to record a pitch test whose scoring was straightforward, but how were we to score a loudness test when it always showed a time-order error? Were we to try to explain to the military that, although the second stimulus may actually have been 0.1 db lower than the first, if the soldier called it lower they should mark him wrong? We tried to get rid of the error by making the transition from the standard level to the variable level instantaneous, but the error was still there. We finally solved the issue by avoiding it: we dropped the stimulus items whose proper scoring was made ambiguous by the time-order error.

Figure 2 shows results obtained with these tests from groups of 50 subjects each. (In the course of this work several hundred subjects were tested.) The crossing points of the psychometric functions show that there is no time-order error for pitch, but that for loudness the error amounts to about 0.3 db. Similar results were obtained when we used the Seashore pure-tone tests. Postman (43) then took up the problem and confirmed our finding that there is typically a time-order error with loudness but not with pitch.

This perhaps is not the place to try to "explain" the time-order error, but I would like to venture an opinion about it. First of all, it is important to note that the error is typically small—a fraction of a j.n.d.—and hence is close to the limiting "noise level" of our measurements. It is therefore not surprising that the published findings resound with the noise of disagreement. Attempts to map the course of the effect as a function of time are particularly discordant, which is good reason for giving up the name "time-error." Its de-

pendence on time is probably fortuitous. As a matter of fact, in an experiment still in progress (by G. M. Shickman) we have obtained the usual "time-error" when the subject lifts the standard and the comparison weight simultaneously, one with each hand.

What the Fechnerian time-order error really seems to depend on is the same basic process that makes the category scales on Class I (prothetic) continua turn out to be concave downward, namely, the asymmetry of sensitivity—the fact that discrimination is better toward the low than toward the high end of the range. The methods commonly used to measure the time-order error are essentially similar to those we use in category scaling. The subject places stimuli in two or more categories like *heavier*, *equal*, or *lighter* and the forces that determine his judgments are the same as in any category scaling procedure. (When the method of adjustment is used with loudness a different type of constant error occurs: the variable tends to be set *higher* than the standard [44], but the sign of the error may reverse when the standard is very loud [61].) The relation between the category scale and the conventional time-order error is perhaps best illustrated by means of Fig. 3, where we see three category scales for lifted weights covering different ranges of stimuli. (For a further discussion of these and other such curves, see [64].) The widest range (Cowdrick [7]) gives the usual category curve, concave downward, and we note that the middle weight in the series is assigned a value *above* the middle of the category scale. When the range is shortened (Guilford and Dingman [25]) the same general features are preserved. The shortest range is the one used by Fernberger (11) to determine a difference limen by the method of single stimuli. This curve, like the others, fails to pass

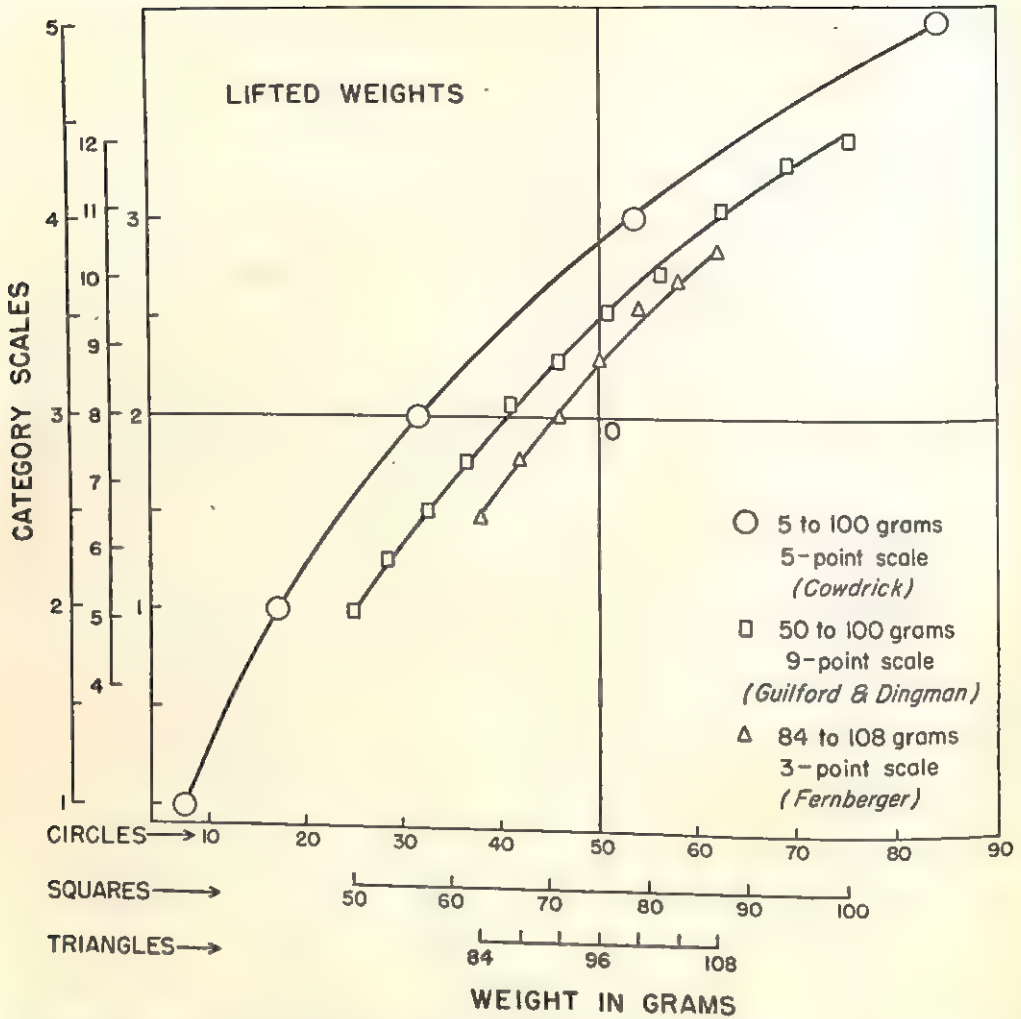


FIG. 3. Category scales for lifted weight. On the middle curve only every other point has been plotted. The end stimuli were designated 4 and 12 but the subject could use other categories if desired. Note that the curves are all concave downward. The fact that they do not pass through the point 0 indicates a "time error."

through the point 0, which is the point defined by the middle of the stimulus range and the middle of the category scale. The distance between the point 0 and the point where the curves cross the horizontal line through the middle of the plot is the time-order error (cf. Pratt [46]). It is as simple as that. By means of the theory of category judgments we can apparently account for Fechner's time-order error without the help of such props as sinking traces, fading images, or assimilation. And it

becomes clear why we should expect to find time-order errors on Class I continua (prothetic) but not on Class II (metathetic).

One further point. Not only is the time-order error produced by other factors than time, but it also has nothing to do with order. Nor is it necessarily an error. The name, like the one given to the Holy Roman Empire, is a misnomer in all particulars. If my thought is correct, that we are dealing here with an effect on category judgments that

derives from an asymmetry based on the relativity of discrimination, a better name for it might be the "category effect." Such a label might also help to distinguish this form of bias from other sources of systematic "error," of which there are many. Not all constant deviations in psychophysical experiments can be blamed on the category effect. To take a single example, the observation of Fullerton and Cattell (13) that the second of two equal light flashes appears fainter than the first is probably due to an altered state of retinal adaptation produced by the first flash. In this instance an adaptation effect can apparently override the category effect and produce what would conventionally be called a positive time error. No doubt there are still other sources of systematic bias in experiments of this sort.

It is true of course that time may play a role in these and other systematic errors, in the sense that the effects may vary with the interval between judgments. But our problem is first to try to discover the basic nature of the asymmetries involved in the various kinds of constant errors and then to see how the forces that produce them may vary with time.

4. *Hysteresis*. Since this word means a lagging behind, as when magnetization lags behind the magnetizing current, it seems a good term to describe what happens when the apparent sense-distances between successive stimuli are judged in different orders. The effect shows up in especially dramatic fashion in experiments on bisection and equisection. In a typical experiment on loudness the subject sits before a row of five keys which he presses to produce the tones (1,000 cps). The levels produced by the two end keys are fixed, 40 db apart, and the subject adjusts the levels controlled by the intermediate keys in order to divide a 40-db interval into four equal-appearing steps in loudness.

Where the subject sets the levels de-

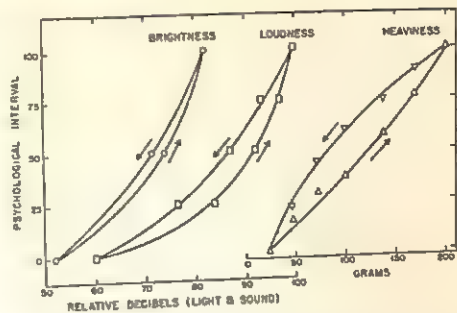


FIG. 4. Hysteresis effects in judgments of intervals on three sensory continua. The arrows indicate the order of stimulus presentation. For brightness and loudness the subject adjusted the middle stimuli to divide the range into two or four equal-appearing intervals. For heaviness the subject lifted three weights and indicated the relative apparent position of the middle weight by means of a slider between two markers on a bar. Five different middle weights were used in ascending and descending order.

pends on whether he listens to the loudnesses in ascending or descending order. I found that in the ascending order he sets the bisecting level some 5 to 8 db higher than in the descending order. In an experiment run to verify this effect, Garner (15) found an average discrepancy of 5.8 db. It is as though the loudness the subject hears lags behind what he should hear as he goes up or down the scale. As a result of this lag the graphs of the functions shown in Fig. 4 exhibit a "hysteresis loop": the ascending path is different from the descending path. Please note, however, that in calling this phenomenon hysteresis I am trying to describe it, not explain it. I am not sure I know how to explain it.

As shown in Fig. 4 the same hysteretic effect occurs in the bisection of brightness intervals and in judgments of the intervals between lifted weights. Partly because the range was shorter (30 db) the brightness bisections show less hysteresis than loudness. Lifted weights, on the other hand, show relatively more. The possibility suggests itself that the magnitude of the hysteresis may vary



inversely with differential sensitivity. In judging the apparent size of the two intervals between three weights, the subjects moved the middle of three pointers on a steel bar to indicate the relative position of the middle weight. Thus we see that hysteresis is independent of whether the subject himself adjusts the stimulus. The same bar and pointers were also used for judgments of loudness, and the resulting hysteresis was essentially similar to that shown in Fig. 4. The data in Fig. 4 are typical examples from an extensive series of unpublished experiments, some of which were run by R. J. Herrnstein.

Hysteresis occurs on the three examples of Class I continua that we have tested. The hypothesis I am suggesting is that it may occur on all continua of this class, but that it probably does not occur on continua of Class II. Evidence for this possibility is suggestive, but not yet conclusive.

Bisections on the continuum of visual position, for example, seem to be free of hysteresis. At least what minor effect has been found is in the opposite direction from that observed on Class I continua. Volkmann kindly undertook to test this matter on the large viewing screen ( $21.6 \times 6.8$  ft.) at Mt. Holyoke College (6). The subject sat 35 feet from the screen and adjusted the position of the middle of three points of light that came on in the sequences right-middle-left and left-middle-right. The end lights were a meter apart. When the 12 observers were instructed to fixate the center of the screen, the two orders produced an average difference in the bisection point of only 0.25 cm. When the instructions were to look at each stimulus as it came on, the discrepancy was 3.6 cm. Since both discrepancies were in the opposite direction from those observed with loudness, brightness, and lifted weights, we may conclude that, whatever constant errors may occur on this Class II continuum,

they apparently constitute a different phenomenon from the Class I hysteresis.

The question of hysteresis in pitch judgments is less easy to answer. We found no evidence for hysteresis in the experiment that led to the construction of the mel scale (67), but the fact is we were not looking for it. In a more direct test of the issue I later had 10 subjects bisect various pitch intervals in ascending and descending order. The experiment was casual and not definitive, and the results were not clear cut. For some subjects hysteresis seemed to be present, for others not. Two subjects happened to have "absolute pitch"—the best cases I happen to have encountered. Each of these two subjects set the middle tone to the same frequency on every trial, regardless of the order of listening. So it seems that at least some people show no hysteresis in their pitch judgments. But this is apparently not true of others.

In a recent laboratory exercise, two of my students, D. D. Greenwood and M. L. Israel, found hysteresis in the equisections of all but one of nine subjects who divided the interval 400–7,000 cps into four equal-appearing intervals.

Incidental evidence for the absence of hysteresis in pitch bisections is found in an experiment by Cohen, Hansel, and Sylvester (5). They were concerned with another problem, but part of their procedure called for the bisection of the pitch distances between 1,000 and 3,000 cps and between 2,000 and 4,000 cps in both ascending and descending order. The bisections were as follows:

For 1,000–3,000 cps, going up: 1,874 cps; going down: 1,838 cps.

For 2,000–4,000 cps, going up: 2,693 cps; going down: 2,808 cps.

The average standard deviation for all settings was 303 cps.

We see that the interval 1,000–3,000 cps shows an insignificant discrepancy (36 cps) and that the larger discrepancy

(115 cps) shown by the interval 2,000–4,000 cps is less than half the standard deviation. It is also important to note that the directions of the discrepancies between the results for ascending and descending orders are opposite in the two parts of the experiment. Incidentally, the points of bisection that are predicted by the mel scale are 1,859 cps and 2,280 cps, which are within 3 per cent of the observed values.

All in all then, the evidence for hysteresis in pitch bisections is ambiguous. The pitch continuum probably means different things to different people, especially to those with and without "absolute pitch." It is in many ways a difficult continuum about which to generalize (see, for example [64]).

Another point concerning bisection deserves mention. On continua of Class I (prothetic) the point of bisection falls consistently below the point predicted by the ratio scale of the subjective magnitude (39, 58, 63). This is true of the combined results from which hysteresis has been eliminated by an averaging of the ascending and descending judgments. It is also clearly evident in the early experiments by Fullerton and Cattell (13) who explored the "force of movement" by the methods of halving, doubling, and bisection. This error in bisection is scarcely surprising when we consider that bisection and category rating-scales have much in common. In both instances the subject tries to equalize intervals, and in both instances he misses the mark in the same direction—presumably because both kinds of judgments are controlled partly by discrimination, and discrimination varies from one end of the continuum to the other.

On continua of Class II (metathetic), on the other hand, the results of bisection agree with those obtained by fractionation and by magnitude estimation. This has been demonstrated for pitch (64, 67) as well as for inclination or

"bearing" (47, 64). Apparently it is only on Class II continua that the construction of a scale of subjective magnitude by the method of bisection leads to a valid outcome, for only there can bisection be shown to agree with direct magnitude estimation.

In our experiments on weight and loudness in which the subject moved a pointer along a bar, the parallel with category rating-scales is even more obvious. In effect, the bar and pointer constitute a continuous rating scale, analogous to the widely used scales on which the rater marks a point on a line with a pencil. The bar and pointer also give results (hysteresis averaged out) that differ from those predicted by the ratio scale of the Class I continuum.

We see then that possibly four functional criteria can be used to divide the perceptual continua into two general classes. Clear examples of each of these classes are easy to find, but this is not to say that all examples one might think of will prove easy to classify. When we try to impose a simple order on the complex fabric of our experience there is an ever-present risk that we will oversimplify what is not simple. Class I (prothetic) constitutes a rather unitary and well behaved group of continua, but Class II (metathetic) includes various disparate aspects of things, many of which do not fall naturally onto unitary monotonic scales, and some of which may not even be orderable. Some continua may turn out to be hybrid mixtures of both Class I and II. Pitch, in fact, may be one of these (64).

In most of what follows we will limit our concern to the prothetic continua (Class I), the "quantitative" aspects of things, for it is here that we can most profitably examine the principles that psychophysics has set itself to devise. The first of these principles concerns the relation between stimulus and sensation.



## THE POWER FUNCTION

There is a growing bundle of evidence to indicate that on prothetic continua the form of the "psychophysical law" is a power function. Most of this evidence, accumulated since 1930, is the product of a reoriented psychophysics that brings "direct" methods to bear on the Fechnerian query regarding the relation between stimulus and sensation. These direct methods lead to ratio scales of perceptual magnitude, and to a first approximation these scales show that the psychological magnitude is a power function of the stimulus magnitude.

The basic principle seems to be that equal stimulus ratios tend to produce equal sensation ratios. Empirical evidence to support this principle has been obtained for at least a dozen perceptual continua, with the aid of several different experimental methods. What the principle entails can be illustrated by examples such as these: In order to make one sound seem half as loud as another, the physical energy must be reduced by about 90 per cent (10 db), and this required reduction is approximately the same regardless of the level from which we start. The same is true of brightness, and the required reduction is of the same order of magnitude as for loudness. In order to make one lifted weight seem half as heavy as another, the original weight must be reduced by about 38 per cent, and this percentage reduction is approximately constant over a wide range of stimuli. The approximate constancy of the percentage reduction corresponding to a given subjective ratio demonstrates that the sensation  $\psi$  is proportional to the stimulus  $S$  raised to a power  $n$ . Thus

$$\psi = kS^n.$$

When we convert this equation to logarithmic form we obtain a linear equation which has a certain practical usefulness, for the function can then be represented by a straight line in log-log

coordinates. The slope of this line corresponds to the exponent  $n$ .

When we know the value of ratio  $r$  between the stimulus judged half and the standard stimulus we can obtain the value of  $n$  by dividing  $\log r$  into  $\log 0.5$ . In general, if  $s$  is the sensation ratio (determined by the method of ratio production) and  $r$  is the corresponding stimulus ratio, then

$$n = \frac{\log s}{\log r}.$$

It is not to be expected, of course, that a simple power function will hold from zero to infinity. What the limits may be must be decided by experiment. We know that some functions relating stimulus to response exhibit discontinuities (e.g., apparent numerosness and flash rate), and we know that our formulas might need eventually to take account of second-order perturbations in the vicinity of the absolute threshold (cf. 8, 58). But our present concern is with first-order approximations to general laws.

## METHODS FOR RATIO SCALING

Before we consider the concrete instances of ratio scales, let us look briefly at the problem of procedure. Methods for constructing ratio scales of subjective magnitude are a relatively new development. It is true that Merkel (31) with his *Methode der doppelten Reize* (1888) tried to find the stimulus that appeared to double the sensation, but Merkel's effort seems to have had little influence on psychophysics. Fullerton and Cattell (13), who used the methods of doubling and halving, fared little better. Even Titchener (72) in his thorough and scholarly Instructor's Manual makes only off-hand allusion to Merkel's method of "doubled stimuli." And of the author of the method he says "we are reminded that he is, by mental constitution, a physicist rather than a psychologist . . . and that his work as a



whole is exceedingly poor in introspective data" (p. 223). It may be that "mental constitution" plays a role in these things, as Titchener suggests. I wonder what he would make of the fact that at least seven different physical laboratories have made important contributions to the development of the ratio scale of subjective loudness—the sone scale—as against perhaps three different psychological laboratories.

The impetus for the development of a method is a problem. Method for its own sake, tempting and fascinating as it may be, often leads to little—except methodology. Physicists and psychologists went to work to refine the procedures for measuring loudness mainly because the substantive outcome was of interest and importance to someone, particularly to acoustical engineers. This is shown by the fact that some of the earliest studies were paid for by commercial companies. Curiously enough the practical problem originated in an obvious failure of Fechner's law. Not long after they had adopted the decibel scale for measuring sound intensities, the engineers noted that equal steps on the logarithmic (decibel) scale do not behave like equal steps, for a level 50 db above threshold does not sound at all like half of 100 db, as Fechner's law implies it should. Since the acoustical engineer must often interpret to his customers the meaning of esoteric acoustical measurements, it soon became clear that a scale was needed on which the numbers would be proportional to how loud things sound to the typical listener. Without the motive of this kind of practical problem to spark the development of method, it is a fair guess that the ratio scales for the fourteen perceptual continua we will examine might never have been constructed. How a subjective ratio scale, once achieved, can be put to work is shown by the way we were able to use the sone scale in the development of a procedure for calcu-

lating the loudness of a complex noise from a spectral analysis of the sound (61).

The methods for constructing ratio scales are still in the process of development, but in one form or another they all require subjects to make quantitative estimates of subjective events. Many authors have screamed that this is nonsensical, meaningless, and impossible (12, 34), but those who follow these methods go ahead and do it anyhow. These direct assessments of sensation seem not so impossible after they have been made.

At present the methods are principally four, but each has subvarieties. We may classify the methods more or less systematically as follows:

1. Ratio estimation
  - a. Direct judgment of ratios
  - b. "Constant-sum"
2. Ratio production
  - a. Fractionation
  - b. Multiplication
3. Magnitude estimation
  - a. Prescribed modulus
  - b. No designated modulus
4. Magnitude production

I am well aware that few activities in psychology exhibit greater terminological confusion than the naming of methods, and that there is probably no effective cure for the burgeoning, ramifying process that occasions the welter of names. The best we can do is take stock from time to time and attempt a little systemization. Having formerly contributed my full share to the confusion, in the foregoing list I have tried for a dose of order.

Partly through my efforts the second class of procedures, which was the first used historically, got the name fractionation (68), because the usual procedure has been to require a subject to set one stimulus to produce a sensation half as great as a given standard. Frac-

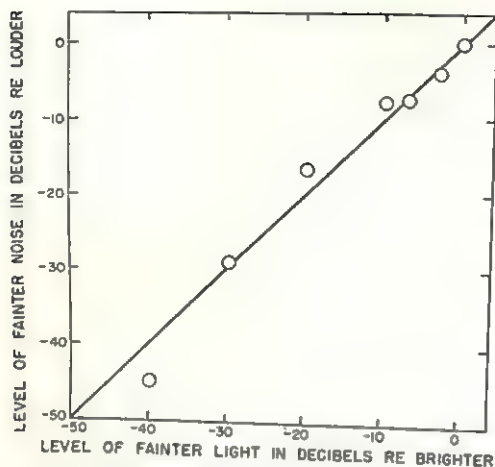


FIG. 5. Results of adjusting a loudness ratio to equal an apparent brightness ratio. The subject viewed a pair of luminous circles of which one was dimmed by an amount shown on the abscissa. He also pressed a pair of keys to produce two different levels of white noise. He adjusted one level (ordinate) to make the loudness ratio seem equal to the brightness ratio. The brighter light was about 99 db re  $10^{-10}$  lambert and the louder noise was about 92 db re 0.0002 microbar.

tions other than a half are also used with consistent results. But fractionation in this sense is only one facet of a more general procedure. The other facet, which we might call *multiplication*, involves the complementary process of requiring the subject to identify or produce a prescribed ratio that is greater than one, e.g., to double, triple, etc., a given standard. This procedure has not been used very much, probably not as much as it should be, for there is abundant evidence (62) to indicate that, when it is made to complement halving, the procedure of doubling can help to balance out certain systematic biases.

The two procedures together may be called *ratio production*. The production of ratios can be carried out by a variety of different techniques. Thus the experimenter may allow the subject to adjust a stimulus to produce a prescribed ratio to a standard, or the experimenter may set the stimulus and ask the sub-

ject whether it meets the prescribed ratio (method of "constant stimuli").

An interesting variant of ratio production involves the fixing of two luminances to define an apparent ratio which the subject is to reproduce by setting two loudnesses to the same apparent ratio. In an experiment by J. C. Stevens in this laboratory the subjects set nearly the same *physical* ratio (decibels) between the intensities of the white noises as the experimenter set between the intensities of the two white surfaces. Figure 5 shows the median settings made by 15 subjects, each of whom twice adjusted the loudness of the second of two noises until the apparent ratio between the noises equaled the apparent ratio between two luminous targets viewed against a dark surround. The subject sat about 4 feet from the targets, which were 4.4 cm. in diameter and separated by about 17.5 cm. The fact that the median settings of the variable noise fall fairly close to the line indicates that the subjective impressions produced by white light and white noise are similar functions of stimulus intensity. The adjustments made to the largest ratio (40 db) were of course quite variable: the semi-interquartile range was about 10 db.<sup>2</sup>

The method of *ratio estimation* is the inverse of ratio production. Instead of prescribing the ratio in advance, the experimenter presents two (or more) stimuli and asks the subject to name the ratio between them. The subject may name the ratio directly, as in the pioneer experiment of Richardson and Ross (49), or he may be constrained to express the ratio by dividing a given number of points between the two stimuli in the manner proposed by Metfessel (32). The constraint involved in the so-called "constant-sum" method has

<sup>2</sup> For the rationale of measuring luminance in decibels see (59).



obvious disadvantages when large ratios are involved.

The method of *magnitude estimation* (60) dispenses with ratios as such and requires the subject to assign numbers to a series of stimuli under the instruction to make the numbers proportional to the apparent magnitudes of the sensations produced. The experimenter may prescribe a modulus by presenting a given stimulus and calling it some particular value, e.g., 10, or he may leave the subject free to pick his own modulus. (Note: if the magnitude estimations give a skewed distribution, as they usually do, it is advisable to calculate medians instead of means.)

The method of magnitude estimation has a logical inverse in *magnitude production*—a possible method that has thus far been mostly ignored. Instead of presenting a series of stimuli in irregular order and asking the subject to judge their apparent magnitudes, the experimenter might name various magnitudes and ask the subject to adjust stimuli to produce proportionate subjective values. Like any given procedure, this method probably has its "bugs" and biases and it might be interesting to find out what some of them amount to.

One thing we know is that in using this method the experimenter must resist whatever impulse he may have to fix or designate the top and bottom of the range, for this will change the problem into one involving category scaling. In some earlier experiments (64) we used this procedure, which we may call category production, to generate a 7-point category scale of loudness. We presented two levels which we designated 1 and 7 and then asked subjects to produce each of the other categories in irregular order. The results are like the usual category judgments obtained on Class I continua: the function is concave downward when plotted against the sone scale.

Magnitude production of a sort has been used in an experiment in which we used a brightness rather than a number to designate a magnitude. We set the luminance of a single target to various levels and asked subjects to adjust a noise to make its loudness appear as great as the brightness of the light. Although this investigation is still in progress (by J. C. Stevens) the results appear to be fairly consistent with what we know about the subjective scales for loudness and brightness. The intensities of the white noise are made roughly proportional to the intensities of the white light, again suggesting that loudness and brightness are similar functions of intensity.

All four of the methods listed above provide the kind of data necessary for the construction of a ratio scale (55). Each method can of course be varied and modified in numerous ways. Not only do the methods need to be altered and adapted to fit the problem at hand, but in any serious effort to establish a definitive scale for a given perceptual continuum it is desirable to ferret out the possible sources of bias by using different methods and by altering relevant parameters. In the present state of the art a valid scale that is representative of the typical subject can scarcely be hoped for from a single try.

The following ratio scales of subjective magnitude have been erected by means of one or more of the general procedures named above. They have not all been given the intensive study and cross-checking they deserve, but our interest here is in their general form rather than in their detailed adequacy. To a first approximation they are all power functions.

#### FOURTEEN RATIO SCALES

Our review of the fourteen examples of ratio scales of subjective magnitude will be brief, because most of the details



TABLE 1

PERCEPTUAL CONTINUA ON WHICH PSYCHOLOGICAL MAGNITUDE  
IS A POWER FUNCTION OF THE STIMULUS

The second column shows the approximate exponent of the power function. Names suggested for the various subjective units are listed in Column 3. The methods used (Column 4) are indicated by numbers that refer to the methods listed elsewhere in this paper.

Continuum	Exponent	Name of Unit	Methods Used	Reference Source
Loudness	0.3	sone	1a, 2a, 2b, 3a, 3b, 4	(58)
Brightness	0.3-0.5	bril	2a, 2b, 3a, 3b	(28), (64)
Visual distance	0.67		2a	(17)
Taste	1.0	gust	2a	(1)
Visual length	1.1	mak	1a, 2a, 3a	(8), (47), (64)
Visual area	0.9-1.15	var	1a, 1b, 2a	(8), (47)
Duration	1.05-1.2	chron	2a, 3a	(19), (50), (64)
Lightness of grays	1.2		3a, 3b	(64)
Finger span	1.3		3a, 3b	*
Numerousness	1.34	numer	2a	(69)
Heaviness	1.45	veg	1a, 1b, 2a, 2b, 3a, 3b	(64)
Auditory flutter	1.7	flut	2a, 2b	(42)
Visual velocity	1.77		2a	(9)
Visual flash rate	2.0		2a	(48)

\* S. S. Stevens and Geraldine Stone. Research in progress.

are available elsewhere. We can help the cause of brevity by means of Table 1 where the significant facts are listed and the sources of more detailed information are tabulated. In Table 1 the continua are ordered by increasing size of the exponent of the power function, and whenever a name for the subjective unit of the ratio scale has been suggested it is recorded. The methods that have been applied to each scale are indicated by numbers that refer to the listing in the previous section. I think all these continua belong to Class I, but this of course may be open to question.

Since Table 1 was prepared, an experiment on smell has been started in this laboratory by T. S. Reese. The early results suggest that the apparent intensity of the odor of benzaldehyde (synthetic almond) is a power function of the concentration in moles per liter of air. The exponent seems to be of the order of 0.2.

It goes without saying that all these scales have not been determined with

equal care and thoroughness. About some of them we know a lot and have reason for confidence in their form; others have received only limited attention. Some of these attributes are easy to scale; others are more resistant to precision. Some of them represent perceptions about which we frequently make judgments of a more or less quantitative sort and about which the subjects in an experiment know, or think they know, so much about the stimulus that the "stimulus error" (3) may be a ready danger; others involve stimuli whose physical parameters are so unfamiliar to the typical subject that his temptation to try to estimate a physical measure of the stimulus, rather than to judge how it appears, is remote and un-compelling. Yet all these ratio scales approximate power functions.

I say approximate because it is certain that not all judgmental continua are strictly power functions of an arbitrary physical parameter of the stimulus. We know, for example, that over

most of the range the loudnesses of a 1,000-cycle tone and a white noise grow with intensity according to similar laws, but at very low intensities the loudness of the white noise grows more rapidly. Hence the loudnesses of noise and tone cannot both follow power functions throughout their entire range, and it is even doubtful that either is strictly a power function at better than a first-order approximation (62).

On this question of the form of a psychophysical law we ought perhaps to try to learn a lesson from Fechner's mistake and forego the temptation to insist on the exactitude of a particular formula. In the realm of human judgment we may aspire to the discovery of first-order invariances, but it would be idle to pretend that precision can be pushed to the level envisaged by Fechner. The judgment of subjective magnitude is inherently a "noisy" phenomenon. When people try to describe a sensation in quantitative terms they face a difficult task, and the factors that affect the outcome are numerous and subtle. Patience and experimental skill can probably clean up part of the variance, but there will always remain irreducible dispersions to set a level below which we sink into uncertainty. In the broader perspective of things the power law stands out as a first-order relation, just as Weber's law describes the first-order relation between magnitude and resolving power. Although we are sometimes more fascinated by second-order departures from first-order relations than we are by first-order generalities have the broader significance. It is of greater moment to know the first-order law for the velocity of falling bodies than it is to prove that a host of variables can produce departures from the law.

Brief comments on some of the scales listed in Table 1 are now in order. We have already discussed loudness in vari-

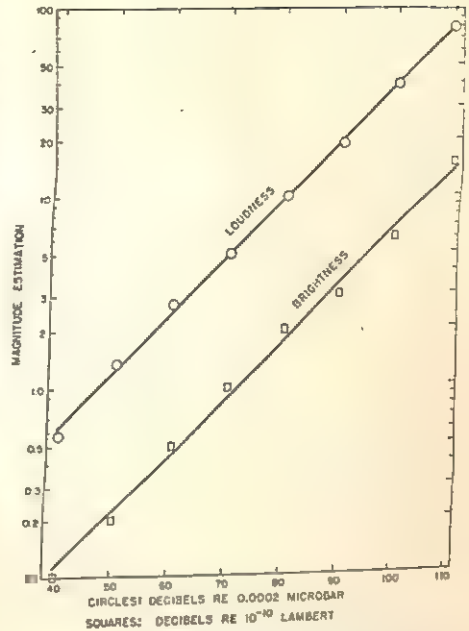


FIG. 6. Median magnitude estimations for loudness and brightness. For loudness each of 32 subjects made two estimates of each level (1,000 cps) in irregular order with no designated standard. Estimates were transformed to a common modulus at the 80-db level. For brightness each of 28 dark-adapted subjects made two estimates of each level in irregular order. The target subtended about  $5^\circ$  and was illuminated for about 3 sec. at each presentation. Once at the beginning of each session the subject was shown the level 70 db (14 subjects) or 80 db (14 subjects) and told to call that level 10. Estimates for all subjects were transformed to a common modulus at 70 db.

ous connections, and a review of most of the work on the problem is available (58). An example of a segment of the loudness function obtained from 32 subjects by the method of magnitude estimation is shown in Fig. 6. It is interesting that the sone scale is, so far as I know, the only subjective ratio scale for which the American Standards Association has appointed a committee to recommend a standard form.

Brightness is still under active study. For ordinary luminous targets seen against a dark surround by a dark-

adapted observer, brightness seems to grow approximately as the 0.3 power of the luminance. Figure 6 shows magnitude estimations for brightness obtained from 28 subjects in an experiment being conducted by J. C. Stevens. For very small targets that approximate point sources the exponent is larger than that determined by the line in Fig. 6, and of course the state of adaptation of the eye and the presence or absence of contrast makes a difference to the brightness function. These two factors increase the size of the exponent, as can be seen in Table 1 where the exponent for lightness of gray papers is listed as 1.2. When grays of different reflectances are viewed under ordinary illumination, light adaptation and contrast both operate to increase the steepness of the brightness function to a different order of magnitude from that observed with luminous targets seen in the dark.

Visual distance refers here to the apparent distance from the observer to a marker on a 200-foot line that stretched away from the observer on a large flat lawn. Two subjects made fractionation judgments. The combined results can be fitted by a  $\frac{2}{3}$  power law as well as, if not better than, by the formula proposed by Gilinsky (17). Specifically, the data in her Fig. 6 fall close to a straight line. The main difficulty with the fit occurs at the shorter distances, where for one observer the apparent half values are nearly equal to the objective half values. But we would expect apparent distance to approximate objective distance when the angle between the line on the grass and the observer's line of regard becomes large. Perhaps if the line went out from the bridge of the observer's nose we would obtain a power function all the way—with an exponent less than one. Of course apparent distance, like apparent size, is specific to the experimental situation and is af-

fectured by cues of various sorts. It is not to be expected, therefore, that a single power function will govern all judgments of distance. As a matter of fact, in a very different set-up involving short distances (4.5 meters and less) Gruber (21) obtained data that fit a power function with an exponent of approximately 1.3.

Taste, curiously enough, gives an exponent of approximately unity when the stimulus measure is the ratio of solute to solvent (1). The functions reported for four different taste substances vary somewhat from one to another, for taste is a troublesome modality to work with, but they all give fairly straight lines in a log-log plot and the slopes hover around the value one. Another experiment (on sweetness [30]) gave only a rather poor approximation to a power function.

Length and duration give subjective ratio scales that are nearly proportional to the physical scales on which we usually measure the stimulus. On balance, however, there is a tendency in both cases for the exponent of the power function to be greater than one, which means that on these continua the magnitude of the subjective impression is a slightly accelerating function of the stimulus magnitude. The accelerating function for duration is shown in Fig. 1.

Apparent area seems to be nearly proportional to physical area, but the shape and size of the area to be judged seem to produce second-order differences. For smaller squares the exponent appears to be greater than for larger squares. In a recent experiment with circles the exponent turned out to be less than one (8).

For continua such as length it is sometimes held that learning plays a major role: we learn what length measures half as long as another, and so it comes to look half as long—which is one way of explaining why the exponent of



the power function is near unity. This may or may not be true. Of course, if by learning we mean that the organism has had a history, we can never hope to experiment on subjects who have not learned. How the organism got to be the way it is will always be an engaging problem, but it is also a worthy object of attention to discover just how the typical subject now behaves, regardless of how he got that way.

The "learning explanation" of these ratio scales seems much less convincing for such continua as loudness and brightness. Despite Ebbinghaus' argument (see [60]) that we learn the meaning of a doubled brightness by illuminating the scene with two lights instead of one, the typical viewer actually requires some nine or ten times the illumination before he reports the brightness doubled. Why is this? And why is it that when the reflectance of gray papers is involved, where it is scarcely conceivable that the typical viewer knows the nature of the physical processes at work in the stimulus, the exponent is greater instead of much less than unity? Sophistication as regards the stimulus seems less than a likely explanation.

It is true, however, that one can learn to estimate the physical magnitude of a stimulus, as a good photographer well knows. He corrects what he perceives by what he knows about light and its relation to subjective impressions. By a similar procedure the acoustical engineer may estimate decibel levels with fair precision. He learns what a given subjective impression means in terms of the indications on a level meter. I have tested at least a half dozen acoustical engineers whose business involves the estimation and measurement of decibel levels. Despite their familiarity with decibels their subjective scales of loudness are not systematically different from those of other listeners. They can tell me, for example, both that one

sound seems about half as loud as another, and also that the one is about 90 db and the other about 100 db.

Let us return now to Table 1. The ratio scale for finger span was obtained as part of a program that is still in progress. Subjects judged the apparent thickness of blocks of wood placed for about a second between the thumb and middle finger. The magnitude estimations obtained show that over a range of from about 0.2 to 6.0 cm the impression of felt thickness grows as a power function of the thickness of the blocks, and that the exponent is about 1.3.

Numerousness refers to the impression one gets from looking at a collection of dots that contains too many to be counted. The exponent of about 1.34 was derived from Taves' experiment in which he used the method of fractionation. I computed the exponent from the mean fraction judged half for all standards above 8 dots. The exponent is equal to  $\log 0.5$  divided by the logarithm of the fraction judged half.

Lifted weights have been scaled by several methods in several laboratories and most experimenters have found close approximations to power functions. The value of 1.45 for the exponent of the veg scale was obtained from pooled results from five different laboratories. For the procedures involved, see (64). It is interesting to note that when Fullerton and Cattell in 1892 had their subjects try to halve and double the apparent force exerted on a dynamometer they came quite close to the values that would be predicted by the veg scale. The two values obtained fell on opposite sides of the predicted value, and the largest discrepancy was about 10 per cent of the predicted value. Although these data are meager they suggest that the apparent force of this type of muscular pull may follow a power function that is like the veg scale.

Auditory flutter refers to the appar-

ent rate at which a white noise is turned on and off. Over a wide range of flutter frequencies (5 to 200 per second) the apparent rate of flutter grows as the 1.7 power of the frequency of stimulus interruption. Pollack used both fractionation and multiplication in determining this function. The one value he tested at a rate below 5 per second suggests that there may be a discontinuity in the function at about this point. Otherwise the data give a close fit to a power function.

Velocity of seen movement gave data that fit a power function so well that, as the authors say, "The goodness of fit may be judged as excellent." The subjects alternately viewed two horizontally moving targets consisting of vertical white and black bands, and they adjusted the velocity of one to make it appear half the other. The standard velocities ranged from 1.8 to 42 mm/sec.

Ekman and Dahlbäck (9) call attention to another interesting fact, namely, that when the subjects were asked to set one velocity *equal* to a standard velocity the variability of the settings was not very different from the variabilities obtained when one velocity was set to appear *half* of a standard. They say, "the degree of uncertainty in fractionation is only a little greater than that in equating stimuli." Hanes (28) had earlier noted the same thing in his experiments on brightness. He says, "Extremely interesting is the fact that the standard deviations around the group means of the equality matches . . . are of the same order of magnitude as the deviations around the means of the  $\frac{1}{2}$  and  $\frac{1}{3}$  estimates at the same brightness levels." I can add for loudness that the variability of halvings and doublings is actually slightly less, in general, than the variability of equality matches made between two noises of substantially different spectra, say, a hiss and a rumble. In experiments involving several thou-

sand loudness matches (61) the standard deviations ranged from 1.5 db, for very similar spectra, to 9.0 db for disparate spectra. Comparable variabilities for 36 subjects who halved and doubled the loudness of a white noise averaged 4.2 db (45). For pure tones, Garner (14) reports variabilities for halving that range from about 3 to 6 db.

Visual flash rate gives the largest exponent yet encountered. For rates between 5 and 10 per second the exponent is about 2.0. Actually the subjective scale for flash rate seems to break up into two segments: one for slow rates where the subject probably judges time elapsed between flashes, and one for faster rates (greater than about 5 per second) where the subject seems to judge rapidity of flash. Except near the transition point, both segments of the scale approximate power functions with slopes decidedly greater than unity.

This completes the list. Let us now look briefly at some of the antecedents of our power law.

#### SOME ANTECEDENTS

The idea that the relation between stimulus and perceptual response might be a power function is no new thing. What is new is the devising of experiments to demonstrate the fact.

The first recorded notion that a power function might be involved appears to be the conclusion reached by Plateau (41) in the 1850's—before the appearance of Fechner's *Elemente* (1860). Plateau reasoned that, since the apparent relations among different shades of gray remain *sensiblement le même* when the general illumination is changed, the *ratios* among the sensations produced by the grays must remain fixed. This, he said, and quite rightly, is more rational than Fechner's view (later revived by Wright [75]) that it is the *differences* that remain constant. Plateau's assumption entails a power function,



Fechner's a logarithmic function. By the present evidence, Plateau was right and Fechner was wrong. Sad to relate, however, Plateau later changed his mind, and for a reason we now can see was irrelevant to the issue: Delboeuf's equisection experiment did not yield a power function. And Fechner, crafty polemicist that he was, disposed of Plateau's power function by the equally irrelevant argument that its champion had repented.

But the power function came back to plague Fechner by another route. Brentano (4) suggested that an increase in a sensation is just noticeable when the increase is a fixed proportion of the original sensation. Brentano did not explore the mathematical consequences of this introduction of a kind of "Weber's law" into the realm of sensation, but Fechner did the mathematics for him (10) and showed that, if Weber's law holds for *both* the sensations and the stimuli, the consequence is a power function. This way of getting to a power function was later argued for by others, particularly Grotenfelt (20), but since there is scarcely any way to prove the basic assumptions involved, the cause was lost under Fechner's massive attacks upon it.

A third approach to the power function was devised by Guilford (22, 23), who called it the "nth-power law." It is based on the Fechnerian premise that j.n.d.'s are subjectively equal and that the scale of sensation can be had by summing them up. Guilford's new twist was to alter Weber's law by making the stimulus increment  $\Delta S$  proportional not to  $S$  but to a power of  $S$ . A Fechnerian integration then yields a power function—which demonstrates that we can sometimes reach a correct conclusion by starting from two wrong assumptions. The assumption that  $\Delta I$  is a power function of  $I$ , with an ex-

ponent lying between 0.5 and 1.0, is not very easily justified by the facts. It is not so easily justified as the more plausible assumption that  $\Delta S = k(S + d)$  where  $k$  and  $d$  are constants. This latter formula fits the data Guilford himself cited better than does the power function, and there is other evidence to support it (33).

But it is Guilford's other assumption that is the greater source of trouble. How serious is the consequence of assuming the equality of j.n.d.'s becomes clear if we begin with any of the power functions (Table 1) whose exponent is greater than 1.0 and follow Guilford's development in the reverse direction. We reach the startling conclusion "that  $\Delta S$  would be a decreasing function of  $S$ , which is unheard of in psychophysical research" (24).

In the light of what we now know, it would appear that of the three historical approaches to the power function Plateau's was based on the most valid assumption. But unfortunately Plateau's premise was only an assumption, and its empirical verification had to wait for nearly a hundred years.

Our position now is more fortunate. The road to the formulation of a general psychophysical law need no longer be strewn with assumptions about j.n.d.'s or contrast steps. The principle I am proposing, that equal sensation ratios are produced by equal stimulus ratios, is merely the summary statement of what we observe on at least a dozen perceptual continua. The power function which this simple principle entails appears to be the first-order relation between stimulus and response. Second-order departures from this law are certain to exist, and their exploration may sometimes prove interesting and instructive, but the wide invariance of the first-order relation deserves to remain the matter of first importance.



## THE SUBJECTIVE SIZE OF THE J.N.D.

The issue raised by Fechner when he assumed that each j.n.d. represents a constant increment in sensation is a stubborn and vexatious problem. It is equivalent to the assumption that measures of resolving power provide equal units that can be used as a measure of magnitude. Applied to Class I (prothetic) continua the assumption is dead wrong, but how can we prove it? Or rather, how can we make the proof persuasive and convincing? Proof has been offered in the past (33, 38, 52) but ingenious counter-arguments have been marshaled against the evidence. Piéron (40) is the objector most persistent and erudite, but others join the chorus. It seems so simple and sensible to take the j.n.d. as a unit of sensation. How can it not be proper?

It is improper simply because it is wrong. Let us consider a concrete case. Both fractionation and magnitude estimation show what is quite obvious to begin with, namely, that the apparent length of a line seen against a homogeneous background is very nearly a linear function of physical length. Two inches look about twice as long as one inch. So we know that subjective length is related to physical length by a power function whose exponent is close to unity. Now what about the j.n.d., the resolving power? We can easily tell visually when one centimeter is added to one centimeter, but we can scarcely tell when a centimeter is added to a meter. So resolving power must be relative, or approximately so. As Münsterberg (37) showed, Weber's law holds fairly well for length. Now we put these two facts together and what happens? We see immediately that in terms of their psychological magnitude, as measured by the scale we first set up by fractionation and magnitude estimation, the j.n.d.'s get larger as we go up

the scale. Q.E.D., the j.n.d. is not constant in subjective size, and counting them off does not yield the magnitude scale.

The principle involved can be clearly seen if we put the problem in more general terms. Let us assume that sensation  $\psi$  is a power function of stimulus magnitude  $S$  and also that Weber's law is true, i.e.,  $\Delta S = kS$ . By the first assumption we may write, disregarding constants that depend only on the choice of units,

$$\psi = S^n.$$

If Weber's law is true, when we count off the j.n.d.'s and record their number  $J$  as a function of the stimulus, we obtain

$$J = \log_c S,$$

where  $c$  is a constant that can serve as the base of the logarithm. We can rewrite this equation as

$$S = c^J$$

and substitute this expression for  $S$  in the first equation, which gives

$$\psi = c^{Jn}.$$

This equation tells us that the psychological magnitude grows as an exponential function of the number of j.n.d.'s  $J$ , which brings us to the interesting conclusion that the scale of sensation is an exponential function of the scale Fechner believed it to be. If we differentiate this equation we get

$$\frac{d\psi}{dJ} = ac^{Jn},$$

where  $a$  is a constant. The equation tells us that instead of being constant as Fechner assumed, the subjective size of the j.n.d. grows as an exponential function of the number of j.n.d.'s above threshold. Simple substitution in this last equation also reveals that the sub-

jective size of the j.n.d. is proportional to the psychological magnitude  $\psi$ .

These equations tell us how it would be if Weber's law were strictly true, but since this law is not strictly true the situation is not quite so simple. The complications, however, are only second order. A superior form of Weber's law is  $\Delta S = k(S + d)$ , which was known to Fechner, Helmholtz and others. By the addition of the small constant  $d$ , this form makes allowance for the inevitable residual "noise level" in the perceptual system, but at the same time it preserves the basic principle that resolving power is relative. If this form is accepted, the foregoing equations and derivations get complicated, but the conclusions they lead to are not substantially altered. It is still true that the subjective size of the j.n.d. grows rapidly as we go up the scale.

Actually, since the value of the constant  $d$  is relatively small, its addition to the formula for Weber's law affects the functions in a significant way only at low stimulus levels. At levels near threshold the general formulas developed above break down, but over the middle and higher stimulus ranges they give a good account of things.

A concrete instance of these general relations is shown graphically in Fig. 7. Miller (33) measured the j.n.d. for white noise by the so-called quantal procedure (65) which provides an estimate of the quantal increment in the stimulus that is just detectable. The results are well described by the relation  $\Delta I = k(I + d)$ . Miller summated the quanta (increments) and obtained the top curve shown in Fig. 7. Except at the low end it is almost a straight line when plotted against decibels. The bottom curve is the sone scale of subjective loudness. The triangles represent the median results obtained by J. C. Stevens and E. Tulving (51) from 70 subjects who made direct magnitude estimations

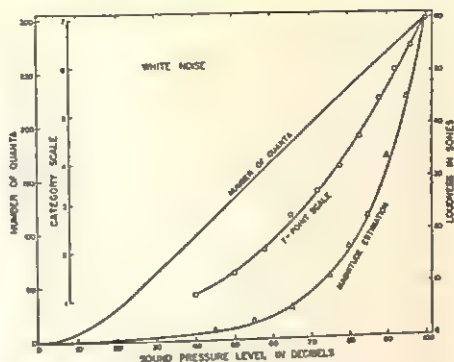


FIG. 7. J.n.d. (quantum) scale, category scale, and ratio scale (sones) for the loudness of white noise. Circles: with the spacing of the levels adjusted to give a "pure" category scale, 8 subjects each made two judgments of each level on a 7-point scale. The range was disclosed at the outset and the order of presentation was different for each subject. Triangles: median judgments by 70 subjects who made magnitude estimations of eight levels presented in irregular order in a classroom, with no designated standard. Estimates were transformed to a common modulus at the 75-db level, which was the first level presented.

of the loudness of different levels of white noise presented in irregular order. In a log-log plot these points fall close to a straight line like that shown in Fig. 6.

The curves in Fig. 7 show how vastly different are the two scales, the j.n.d. (or quantum) scale and the loudness scale. In this respect they confirm what we have already seen in Fig. 1. They also show that when we use the loudness scale to measure the subjective size of successive j.n.d.'s, their size grows rapidly as we go up the scale. Except near the low end, their growth approximates an exponential function. Actually, over the low end of the scale where the j.n.d. scale is curved there is a superficial similarity between the j.n.d. scale and the loudness scale, but over the major portion of the audible range the divergence between the two scales is large and obvious.

The middle curve in Fig. 7 shows an



example of the function obtained when the loudness of a white noise is judged on a scale from 1 to 7 (data from Stevens and Galanter [64], Fig. 9A). The spacing of the stimuli had been adjusted to give a "pure" category scale. Just as in Fig. 1, we see that the form of the category scale is intermediate between that of the j.n.d. scale and the scale of subjective magnitude. If plotted against the sone scale the category scale would be concave downward, as are all category scales on prothetic continua.

It has occasionally been asserted that some scales of sensory magnitude agree with the scale obtained by counting off j.n.d.'s (74). Hanes' (27) statement that this is true for brightness is a recent example, which has led, incidentally, to a certain amount of confusion. In a subsequent study Hanes (28) extended the range of his measurements far enough for us to see clearly that the scale of subjective brightness (in brils) departs from the j.n.d. scale in the expected way. Perhaps Hanes should have stated this explicitly.

I have summed the j.n.d.'s for brightness obtained by three different experimenters (2, 18, 35) and have compared the results with Hanes' bril scale. The pictures presented are essentially similar to that shown for loudness in Fig. 7, except that the curved tail at the lower end of the j.n.d. scale tends to be longer. The basic fact is that over the middle and upper ranges of intensity Weber's law holds fairly well for brightness, but the bril scale is a power function. Over the range of the fainter luminances, Weber's law breaks down to a degree sufficient to produce the apparent agreement Hanes spoke of. But it is quite wrong to jump to the general conclusion that j.n.d.'s add up to the bril scale.

We have already noted, in the discussion of Guilford's "nth-power law," that in order for the j.n.d.'s to add up to give the magnitude function,  $\Delta S$

cannot in general be an increasing function of  $S$ . For most perceptual continua it must be a decreasing function of  $S$ . But this is "unheard of in psychophysical research." It would mean, for example, that it would be easier to tell when a centimeter was added to a meter than when a centimeter was added to a centimeter. Resolving power simply does not work that way.

Suppose now for a moment we risk Titchener's displeasure and try to think like a physicist. Physics also faces the problem of resolving power and in most physical measurements resolving power is relative—the accuracy of measurement is proportional to magnitude. This fact is expressed by saying that length can be measured to 1 part in  $10^5$ , that resistance can be measured to 1 part in  $10^6$ , or that time intervals can be measured to 1 part in  $10^8$ . In this sense Weber's law has its analogue in physics. Now, it is conceivable that someone might propose that we give up our ordinary ways of measuring length, say, and devise instead a scale whose unit is the resolving power of our best procedures. We would then be doing what Fechner proposed. In terms of our present units, the new units of length would grow increasingly larger as we go up the scale, and they would do so for the same reason that the j.n.d.'s grow larger as we go up the psychological scale on a prothetic continuum. The laws of physics could be rewritten in terms of the new scales of units based on resolving power, but it would probably strike the physicist as less than sensible to try to do it—and not simply because it would be hard work. For reasons that he could readily justify, the physicist wants to start with magnitudes and not with resolving power. What I am suggesting is that there is sufficient parallel between physics and psychophysics to justify the same approach in both fields. In neither discipline should we try to



make resolving powers, variabilities, confusions, j.n.d.'s, discriminial dispersions or any such entities serve as units of magnitude on continua that behave like the prothetic continua of Class I.

### SCALING THEORY

The same counsel recommends itself to those of us who devise scaling procedures for other than perceptual variables. As Gulliksen (26) complains, those who worry about psychophysical measurement often ignore the work on nonsensory scaling, and vice versa. Titchener (71) also once complained that psychologists have ignored the advances in subjective scaling achieved by the meteorologists. Insularity is probably inevitable in an age of specialization, but it has its dangers. One of the dangers is that the laws of human judgment developed under simple conditions, where we can better see what is going on, may be ignored when we work in more complicated realms where discernment is more difficult.

A large segment of psychological scaling theory has been built around models that make scale "units" out of some measure of "discriminal dispersion," but this is precisely the kind of unit that is demonstrably *not* invariant on prothetic continua. Variabilities and confusions provide a tempting starting point for psychological measurement, chiefly because dispersion among our judgments is something we always have with us, but the transforming of unreliability, inconstancy, or confusion into units of measure somehow needs more than assumptions to back it up. Our fundamental scales can, I think, stand on firmer ground than the mere scatter of our data.

It is recognized, of course, that measures of confusion often have a usefulness of their own, and that for certain purposes it is desirable to construct

"equal discriminability" scales (16). The engineering of schemes for optimizing the coding of information on cathode-ray tubes presents an interesting example of how such measures may be used in the prediction and control of confusability (36). But our concern here is with a different problem, namely, whether measures of confusion can be used as units of magnitude.

The use of dispersion as a tool for magnitude scaling requires assumptions regarding its behavior. How does dispersion (in psychological units) vary over the psychological continuum in question? A variety of assumptions regarding the behavior of variability have been elaborated, but so far as I know, none of them come close to admitting the drastic growth in dispersion that we find to be true for Class I (prothetic) continua. Here, we recall, the dispersion (in subjective units) approaches an exponential function of the number of j.n.d.'s above threshold or, alternatively stated, it grows almost in direct proportion to the psychological magnitude. The assumption made in practice is usually that dispersion (in psychological units) is constant, or nearly so. This is like Fechner's assumption that j.n.d.'s are equal in subjective size. Direct experiments can sometimes verify this assumption on continua of Class II (metathetic), but on the "quantitative" continua of Class I they show that the assumption fails by a wide margin.

On continua that behave like Class I we would be closer to right if we began with the assumption that discriminial dispersion is not constant but is proportional to the psychological magnitude in question. When the psychological magnitude is a power function of the stimulus, this assumption is equivalent to saying that psychological values separated by equal units of dispersion on the stimulus scale stand in a constant ratio to each other. If

we were to start from this assumption, could we then proceed to construct a useful scale of the psychological magnitude from measurements made of discriminational dispersion? In principle we might erect a new type of scale, but we could not construct an interval or a ratio scale of the kinds we ordinarily employ. While working on the theory of scales (54) back in the late 1930's I had a hard time convincing myself that if we know concerning a series of values only that  $a/b = b/c = c/d = \dots$  we are powerless to proceed further toward what I have called a ratio scale of measurement. But the fact is that the equating of ratios gets us no further unless (a) we can also equate intervals, or (b) we can determine the numerical value of the ratio (as we do under the ratio methods reviewed above). An example of how we might use equated ratios *plus* equated intervals to construct a ratio scale is discussed elsewhere (55, p. 24).

But let us pursue for a moment the problem of what we might do when we have equated a set of ratios:  $a/b = b/c = c/d$ . . . . We can assume that we have an operation for ordering these values and that  $a < b < c < d$ . . . . The problem then is, how may we assign numerical values to this series? As is true when we have a series of equated intervals,  $a - b = b - c = c - d = \dots$ , we can pick any two values arbitrarily (subject only to the ordinal requirement), and having done so, all the other values are determined. Thus for the ratio series, if  $a = 2$  and  $b = 6$  then  $c = 18$ ,  $d = 54$ , etc. The next question is, what transformations are permitted on this scale? We can show that any value  $x$  may be replaced by  $x'$  provided  $x' = kx^n$  where  $k$  and  $n$  are constants that may take on any positive values. In other words we can always transform the scale values by a power function. For example, if we square all the values, the ratios remain equal.

A scale of this kind may be mathematically interesting, but, like many mathematical models, it has thus far proved empirically useless. At least, so far as I am aware, no uses for it have yet been found. In a formal sense there are certain parallels between a scale of equal but indeterminate ratios and a scale of equal but indeterminate intervals. On both, for example, any two values can be selected arbitrarily, and the scale values of both can be transformed by an equation involving two arbitrary constants. But the scale of equal intervals has great utility, as witness the Fahrenheit and Celsius scales of temperature, and the scale of calendar time. And differences on interval scales can be measured on ratio scales, as when we measure intervals of duration in seconds. But what can we do with equated ratios when the value of the ratio is indeterminate?

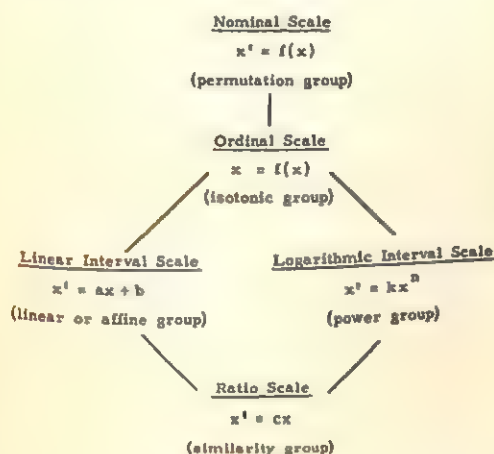
An obvious suggestion is that we might convert to logarithms and express the equated ratios as  $\log a - \log b = \log b - \log c = \log c - \log d$ , etc. If we then restrict the values of  $a$ ,  $b$ ,  $c$ , . . . to positive numbers, we can set up an interval scale in logarithms. There is no a priori reason why we could not put this scale to work in a manner analogous to the workings of our linear interval scales. The linear and the logarithmic scales have in common the fact that the choice of the zero point is arbitrary. It might be interesting to explore the consequences of this procedure and perhaps try to answer the question why science has never made use of these logarithmic interval scales based on equated ratios.

Another question arises. If we were to develop such scales, what would we call them? When I originally proposed (54) the names *nominal*, *ordinal*, *interval*, and *ratio* for the four classes of scales commonly used, I worried some about the scale based on equated ratios,



but I dismissed it as a purely academic question, for there appeared to be no examples of its use. If uses for it were to develop, we would like, perhaps, to call this class of scales *ratio* scales, for they are based on an empirical operation for equating ratios. But since the term ratio scale has now gained wide currency as the name for the scales that admit only a similarity transformation (multiplication by a constant) it may be better to call the scale based on equated ratios a *logarithmic interval* scale.

With this class of scales included, the hierarchy of scales, together with the group of transformations permitted by each scale, would take the form shown in the diagram.



As this diagram suggests, a ratio scale is possible whenever empirical operations are available to create both types of interval scales. If we can determine both equal differences and equal ratios we can eliminate the additive constant  $b$  and the exponent  $n$ , and we are left with only a multiplying constant  $c$ . This defines a scale with a zero point—what we call a ratio scale, or, if we prefer, a “zero-point” scale.

Returning now to the assumptions made about discriminial dispersion we note that, in terms of the scales listed

in the diagram, the assumption most often made is that the processing of variability will yield a linear rather than a logarithmic interval scale. It is not only in connection with the method of pair comparisons, which Thurstone called the “best of all the psychophysical methods” (70), that the wrong assumption of constant dispersion (in psychological units) often gets made. This assumption is either latent or explicit in many subtler procedures, such as the methods of “successive intervals,” “graded dichotomies,” “successive categories,” etc. (see 23). The foregoing arguments suggest that whenever these methods are applied to prothetic continua, continua on which the j.n.d. (dispersion) tends to vary directly with magnitude, the resulting scales will turn out to be nonlinear. It seems obvious that none of the scales listed in Table 1 could be constructed by the method of pair comparisons, or by any of its several cousins. The laws of judgment derived from direct experiments on perceptual continua make it plain that, unless it can be demonstrated that the judgmental continuum behaves like those of Class II (metathetic), any scaling procedure that is geared to the assumption that equally often noticed differences are equal is defective.

Fortunately, this does not mean that no scaling is possible. We still have other devices, such as the direct methods for ratio scaling. Is there a genuine, nontrivial, substantive problem calling for a psychological scale for which these methods, or some variation on them, cannot produce an acceptable answer? Or is it merely that we like to assume, with Fechner, that indirection is the best path to the goal? Contrary to a common assumption, the use of the direct methods does not require knowledge of an underlying, measurable physical continuum (48). Only a nominal scale (55) is required at the stimulus level,



i.e., the stimuli must be identifiable by the experimenter. For example, when the observer judges the apparent ratio between two stimuli the experimenter needs to know which two stimuli were involved, but he does not need to know their values on any other scale. I know from experience that it is more comfortable to take it for granted that a direct method is impossible than it is to try to work one out. And it is still more onerous to test and refine the direct methods and to purge them of constraints and biases. The problem of the scaling of psychological continua is full of booby traps, but so is the laboratory measurement of a physical continuum like, for instance, electrical inductance or mechanical impedence.

Perhaps one of our professional idiosyncracies is that in psychology we are sometimes more enamored of models and methods than we are of problems. One occasionally gets the impression that there are more people with a method who are looking for a problem to use it on than there are searchers with a problem looking for a method. It is admittedly more entertaining to make a formal model, complete with assumptions, postulates, and derivations, than it is to grub through the empirical tangle of an experimental issue. But if something empirically useful is to issue from a model, something empirically known must be put into it. The gratuitous assumption that some measure of dispersion, sensitivity, or resolving power can be used as the unit of a scale of psychological magnitude does not meet this requirement.

#### SUMMARY

This paper has rambled far enough afield at this point and it is time to try to restate the issues in briefer compass. The main points are these.

Two general classes of perceptual con-

tinua can be distinguished by means of four functional criteria. On Class I or "quantitative" continua the j.n.d. increases in subjective size as psychological magnitude increases, category rating-scales are concave downward when plotted against psychological magnitude, comparative judgments exhibit a time-order error (a "category effect"), and equisection experiments exhibit hysteresis. On Class II or "qualitative" continua these four effects are apparently absent. Class I, called *prothetic*, includes those continua on which discrimination is mediated by an additive mechanism at the physiological level; Class II, called *metathetic*, includes those mediated by a substitutive mechanism.

On Class I (*prothetic*) continua the use of one or more of four kinds of direct methods for constructing ratio scales reveals that equal stimulus ratios tend to produce equal subjective ratios. Hence, to a first-order approximation the "psychophysical law" relating stimulus and response is a power function. The exponent, as measured on fourteen different continua, varies from about 0.3 for loudness to about 2.0 for visual flash rate. A few workers in the past have conjectured this power function, even starting sometimes from wrong assumptions. Only lately, however, is this law becoming securely anchored in experiment.

Fechner's logarithmic law is not found in experiment for the simple reason that resolving power (the j.n.d.) is not constant in psychological units, but is roughly proportional to psychological magnitude. For this reason, all procedures of Fechnerian extraction, like the method of pair comparisons and its related techniques, which seek to build scales out of "unitized" measures of dispersion, are not proper methods for scaling magnitudes that behave like Class I or *prothetic* continua. When-

ever psychological scales are called for, direct ratio scaling methods should probably be tried.

## REFERENCES

1. BEEBE-CENTER, J. G., & WADDELL, D. A general psychological scale of taste. *J. Psychol.*, 1948, 26, 517-524.
2. BLACKWELL, H. R. Contrast thresholds of the human eye. *J. opt. Soc. Amer.*, 1946, 36, 624-643.
3. BORING, E. G. The stimulus-error. *Amer. J. Psychol.*, 1921, 32, 449-471.
4. BRENTANO, F. *Psychologie vom empirischen Standpunkte*. Th. 1. Leipzig: Dunker & Humblot, 1874.
5. COHEN, J., HANSEL, C. E. M., & SYLVESTER, J. D. Interdependence of temporal and auditory judgments. *Nature*, 1954, 174, 642.
6. CORBIN, H. H., REESE, E. P., REESE, T. W., & VOLKMANN, J. Experiments on visual discrimination, 1952-55. Psychophysical Research Unit, Mt. Holyoke College, AFCRC TR No. 56-52, April 1956.
7. COWDRICK, M. The Weber-Fechner law and Sanford's weight experiment. *Amer. J. Psychol.*, 1917, 28, 585-588.
8. EKMAN, G., BERGSTROM, B., & KÜNNAPAS, T. M. Comparison between two psychophysical scaling methods. Report from the Psychological Laboratory, Univer. of Stockholm, No. 37, August 1956.
9. EKMAN, G., & DAHLBÄCK, B. A subjective scale of velocity. Report from the Psychological Laboratory, Univer. of Stockholm, No. 31, February 1956.
10. FECHNER, G. T. *In Sachen der Psychophysik*. Leipzig, 1877.
11. FERNBERGER, S. W. On absolute and relative judgments in lifted weight experiments. *Amer. J. Psychol.*, 1931, 43, 560-578.
12. Final Report. *Advancement of Science*, 1940, No. 2, 331-349.
13. FULLERTON, G. S., & CATTELL, J. McK. *On the perception of small differences*. Philadelphia: Univer. of Pennsylvania Press, 1892.
14. GARNER, W. R. Some statistical aspects of half-loudness judgments. *J. acoust. Soc. Amer.*, 1952, 24, 153-157.
15. GARNER, W. R. A technique and a scale for loudness measurement. *J. acoust. Soc. Amer.*, 1954, 26, 73-88.
16. GARNER, W. R., & HAKE, H. W. The amount of information in absolute judgments. *Psychol. Rev.*, 1951, 58, 446-459.
17. GILINSKY, ALBERTA S. Perceived size and distance in visual space. *Psychol. Rev.*, 1951, 58, 460-482.
18. GRAHAM, C. H., & BARTLETT, N. R. The relation of size of stimulus and intensity in the human eye. III. The influence of area on foveal intensity discrimination. *J. exp. Psychol.*, 1940, 27, 149-159.
19. GREGG, L. W. Fractionation of temporal intervals. *J. exp. Psychol.*, 1951, 42, 307-312.
20. GROTFENFELT, A. *Das Webersche Gesetz und die psychische Relativität*. Helsingfors: Frenckell, 1888.
21. GRUBER, H. E. The relation between perceived size and perceived distance. *Amer. J. Psychol.*, 1954, 67, 411-426.
22. GUILFORD, J. P. A generalized psychophysical law. *Psychol. Rev.*, 1932, 39, 73-85.
23. GUILFORD, J. P. *Psychometric methods*. (2nd ed.) New York: McGraw-Hill, 1954.
24. GUILFORD, J. P., & DINGMAN, H. F. A validation study of ratio-judgment methods. *Amer. J. Psychol.*, 1954, 67, 395-410.
25. GUILFORD, J. P., & DINGMAN, H. F. A modification of the method of equal-appearing intervals. *Amer. J. Psychol.*, 1955, 68, 450-454.
26. GULLIKSEN, H. Paired comparisons and the logic of measurement. *Psychol. Rev.*, 1946, 53, 199-213.
27. HANES, R. M. A subjective brightness scale. *J. exp. Psychol.*, 1949, 39, 438-452.
28. HANES, R. M. The construction of subjective brightness scales from fractionation data: a validation. *J. exp. Psychol.*, 1949, 39, 719-728.
29. KARLIN, J. E., & STEVENS, S. S. Auditory tests for the ability to discriminate the pitch and loudness of noises. Report from the Psycho-Acoustic Laboratory, Harvard Univer., OSRD No. 5294 (PB 39576), 1 August 1945.
30. MACLEOD, S. A construction and attempted validation of sensory sweetness scales. *J. exp. Psychol.*, 1952, 44, 316-323.
31. MERKEL, J. Die Abhängigkeit zwischen Reiz und Empfindung. *Phil. Stud.*, 1888, 4, 541-594; 1889, 5, 245-291; 1889, 5, 499-557.
32. METFESSEL, M. F. A proposal for quan-

- titative reporting of comparative judgments. *J. Psychol.*, 1947, 24, 229-235.
33. MILLER, G. A. Sensitivity to changes in the intensity of white noise and its relation to masking and loudness. *J. acoust. Soc. Amer.*, 1947, 19, 609-619.
  34. MOON, P. *The scientific bases of illuminating engineering*. New York: McGraw-Hill, 1936.
  35. MUELLER, C. G. Frequency of seeing functions for intensity discrimination at various levels of adapting intensity. *J. gen. Physiol.*, 1951, 34, 463-474.
  36. MULLER, P. F., et al. Symbolic coding of information on cathode ray tubes and similar displays. *WADC Tech. Rep. 55-375*. Wright Air Development Center, Ohio, 1955.
  37. MÜNSTERBERG, H. A psychometric investigation of the psychophysical law. *Psychol. Rev.*, 1894, 1, 45-51.
  38. NEWMAN, E. B. The validity of the just noticeable difference as a unit of psychological magnitude. *Trans. Kansas Acad. Sci.*, 1933, 36, 172-175.
  39. NEWMAN, E. B., STEVENS, S. S., & VOLKMANN, J. On the method of bisection and its relation to a loudness scale. *Amer. J. Psychol.*, 1937, 49, 134-137.
  40. PIÉRON, H. *The sensations, their functions, processes and mechanisms*. New Haven: Yale Univer. Press, 1952.
  41. PLATEAU, M. H. Sur la mesure des sensations physique, et sur la loi qui lie l'intensité de ces sensations à l'intensité de la cause excitante. *Bull. de l'Acad. Roy. Belg.*, 1872, 33, 376-388.
  42. POLLACK, I. Auditory flutter. *Amer. J. Psychol.*, 1952, 65, 544-554.
  43. POSTMAN, L. The time-error in auditory perception. *Amer. J. Psychol.*, 1946, 59, 193-219.
  44. POSTMAN, L. Time-error as a function of the method of experimentation. *Amer. J. Psychol.*, 1947, 60, 101-108.
  45. POULTON, E. C., & STEVENS, S. S. On the halving and doubling of the loudness of white noise. *J. acoust. Soc. Amer.*, 1955, 27, 329-331.
  46. PRATT, C. C. Time-error in the method of single stimuli. *J. exp. Psychol.*, 1933, 16, 798-814.
  47. REESE, E. P. (Ed.) *Psychophysical research summary report, 1946-1952*. Psychophysical Research Unit, Mt. Holyoke College. USN Spec. Dev. Ctr. *Tech. Rep. No. SDC-131-1-5*, 1953.
  48. REESE, T. W. The application of the theory of physical measurement to the measurement of psychological magnitudes, with three experimental examples. *Psychol. Monogr.*, 1943, 55, No. 3 (Whole No. 251).
  49. RICHARDSON, L. F., & ROSS, J. S. Loudness and telephone current. *J. gen. Psychol.*, 1930, 3, 288-306.
  50. ROSS, S., & KATCHMAR, L. The construction of a magnitude function for short time intervals. *Amer. J. Psychol.*, 1951, 64, 397-401.
  51. STEVENS, J. C., & TULVING, E. Loudness estimation in a group situation. *Amer. J. Psychol.* (in press).
  52. STEVENS, S. S. A scale for the measurement of a psychological magnitude loudness. *Psychol. Rev.*, 1936, 43, 405-416.
  53. STEVENS, S. S. On the problem of scales for the measurement of psychological magnitudes. *J. Unif. Sci.*, 1939, 9, 94-99.
  54. STEVENS, S. S. On the theory of scales of measurement. *Science*, 1946, 103, 677-680.
  55. STEVENS, S. S. Mathematics, measurement, and psychophysics. In S. S. Stevens (Ed.), *Handbook of experimental psychology*. New York: Wiley, 1951.
  56. STEVENS, S. S. Pitch discrimination, mels, and Kock's contention. *J. acoust. Soc. Amer.*, 1954, 26, 1075-1077.
  57. STEVENS, S. S. On the averaging of data. *Science*, 1955, 121, 113-116.
  58. STEVENS, S. S. The measurement of loudness. *J. acoust. Soc. Amer.*, 1955, 27, 815-829.
  59. STEVENS, S. S. Decibels of light and sound. *Physics Today*, 1955, 8 (10), 12-17.
  60. STEVENS, S. S. The direct estimation of sensory magnitudes: loudness. *Amer. J. Psychol.*, 1956, 69, 1-25.
  61. STEVENS, S. S. The calculation of the loudness of complex noise. *J. acoust. Soc. Amer.*, 1956, 28, 807-832.
  62. STEVENS, S. S. Concerning the form of the loudness function. *J. acoust. Soc. Amer.* (in press).
  63. STEVENS, S. S., & DAVIS, H. *Hearing: its psychology and physiology*. New York: Wiley, 1938.
  64. STEVENS, S. S., & GALANTER, E. H. Ratio scales and category scales for a dozen perceptual continua. Report from the Psycho-Acoustic Laboratory, Harvard Univer., PNR-186, August 15, 1956. (Submitted to *J. exp. Psychol.*)
  65. STEVENS, S. S., MORGAN, C. T., & VOLK-



- MANN, J. Theory of the neural quantum in the discrimination of loudness and pitch. *Amer. J. Psychol.*, 1941, 54, 315-355.
66. STEVENS, S. S., & POULTON, E. C. The estimation of loudness by unpracticed observers. *J. exp. Psychol.*, 1956, 51, 71-78.
67. STEVENS, S. S., & VOLKMANN, J. The relation of pitch to frequency: a revised scale. *Amer. J. Psychol.*, 1940, 53, 329-353.
68. STEVENS, S. S., VOLKMANN, J., & NEWMAN, E. B. A scale for the measurement of the psychological magnitude pitch. *J. acoust. Soc. Amer.*, 1937, 8, 185-190.
69. TAVES, E. H. Two mechanisms for the perception of visual numerosness. *Arch. Psychol.*, N. Y., 1941, No. 265.
70. THURSTONE, L. L. Fechner's law and the method of equal appearing intervals. *J. exp. Psychol.*, 1929, 12, 214-224.
71. TITCHENER, E. B. The psychophysics of climate. *Amer. J. Psychol.*, 1909, 20, 1-14.
72. TITCHENER, E. B. *Experimental psychology*. Vol. II, Part II (Instructor's Manual). New York: Macmillan, 1923.
73. WOODROW, H. Time perception. In S. S. Stevens (Ed.), *Handbook of experimental psychology*. New York: Wiley, 1951. Chap. 32.
74. WOODWORTH, R. S., & SCHLOSBERG, H. *Experimental psychology*. New York: Holt, 1954.
75. WRIGHT, W. D. *Researches on normal and defective colour vision*. St. Louis: Mosby, 1947.

(Received for early publication November 2, 1956)

## THEORY OF SELECTIVE LEARNING WITH PROBABLE REINFORCEMENTS

FRANK RESTLE<sup>1</sup>

*Center for Advanced Study in the Behavioral Sciences<sup>2</sup>*

This paper presents a quantitative description of some simple selective-learning data. The descriptive system is shown to subsume the writer's theory of discrimination learning (11), and to relate closely to Estes' statistical theory (5).

Recent mathematical formulations of learning (2, 3, 4, 5, 6, 11) describe the stimulus situation as a set of elements, each of which is conditioned to (tends to evoke) exactly one response at a given time. During learning, if a certain response  $A_1$  is reinforced, a cue may switch and become newly conditioned to  $A_1$ . The probability of such a change is the rate-of-learning parameter. Estes (4, 5, 7) and Bush and Mosteller (2, 3) have shown that such a model has many consequences for learning with random reinforcement schedules, correctly predicting asymptotes in guessing experiments (3, 5, 7). However, certain discrepancies have appeared when theoretical expectations are compared with performance on early trials. Estes<sup>3</sup> has listed some problems which, if solved, would contribute to a workable descriptive theory of elementary learning. One of these is the tendency of the parameter representing learning rate to vary systematically with rein-

forcement probabilities. In general, learning is more rapid when the probabilities of reinforcement are more extreme. A second problem is that subjects can master discrimination problems to perfection even in the presence of "overlap" stimuli which are not differentially reinforced. The writer's discrimination theory (11) has some success in resolving the stimulus-overlap problem by assuming that irrelevant cues are adapted. This paper extends this discrimination theory to cases of random reinforcement, using Estes' equations for probability learning, but, in addition, offering a descriptive account of how learning rate may depend on reinforcement probability.

The major points to be made in this paper are that (1) the stimulus situation in two-choice learning experiments is represented by a set of discriminable aspects called "cues"; (2) a cue may be "conditioned" to either response; (3) a cue may be "adapted" and rendered nonfunctional during learning; (4) the probability of a response is the proportion of the unadapted cues conditioned to it; (5) the degree to which a cue is differentially reinforced determines its "validity"; and that (6) the rates of conditioning and adaptation depend on the validities of the cues available.

When these assumptions are stated quantitatively, the theory is shown to have three characteristics: first, under certain limiting experimental conditions it leads to a process of probability learning similar to that in Estes' theory, and in particular with the same

<sup>1</sup> Appreciation is due Dr. W. K. Estes for his advice and criticism, and for lending pre-publication drafts of papers. The writer has benefited from extended seminars and conversations with L. B. Wyckoff, P. Suppes, and D. H. Howes.

<sup>2</sup> Now at Michigan State University.

<sup>3</sup> See Estes, W. K., The statistical approach to learning theory (draft), prepared for the report of APA "Project A," *Analysis of psychological science*.

asymptotes; second, when applied to discrimination learning it has the same consequences as the writer's earlier discrimination-learning theory (11), with the same way of handling stimulus overlap; and third, it describes the connection between reinforcement probabilities and learning rates in the simplest probability-learning experiments.

## THEORY

### *The Set of Cues*

We think of the stimulus situation as composed of a set of discriminable aspects, called "cues,"  $k, k', k''$ . The set of cues is called  $K$  and the number of cues is  $N$ . The proportion of cues belonging to any manipulated aspect of the situation is measured from behavior, and the number of cues,  $N$ , is for most purposes arbitrary.

### *Conditioning of Cues*

Following Estes (4), we assume that a cue is conditioned to one or the other response alternative at any time, on an all-or-none basis. The probability that cue  $k$  is conditioned to response  $A_1$  at trial  $n$  is called  $F(k, n)$ . If a cue is conditioned to  $A_2$  and then  $A_1$  is reinforced, it may switch over and become conditioned to  $A_1$ . The probability of such a switch is a constant called  $\theta$ , the "rate-of-learning" parameter. On this assumption, we get the following equations of change of  $F(k, n)$ :

i) If  $A_1$  is reinforced on trial  $n$ ,

$$F(k, n+1) = F(k, n)(1-\theta) + \theta. \quad (1)$$

ii) If  $A_2$  is reinforced on trial  $n$ ,

$$F(k, n+1) = F(k, n)(1-\theta). \quad (2)$$

If  $A_1$  is reinforced on some fixed proportion  $\pi$  of the trials and  $A_2$  is reinforced on all other trials, the mean

equation is

$$F(k, n+1) = F(k, n)(1-\theta) + \pi\theta \quad (3)$$

which has the solution

$$F(k, n_0 + n) = \pi - [\pi - F(k, n_0)](1-\theta)^{n-1}. \quad (4)$$

Equation 4 is the conditioning equation which will be used in the applications discussed in this paper.

### *Adaptation of Cues*

During learning, a cue may become "adapted" (11) and lose its effect on response. Different cues have different probabilities of becoming adapted. If cue  $k$  is not adapted by the beginning of trial  $n$ , the probability that it will be adapted by trial  $n+1$  is  $\phi(k)$ . This consideration gives us an equation for  $a(k, n)$ , the probability that cue  $k$  is adapted at the beginning of trial  $n$ , as follows:

$$a(k, n+1) = a(k, n) \times [1 - \phi(k)] + \phi(k), \quad (5)$$

which has the solution

$$a(k, n_0 + n) = 1 - [1 - a(k, n_0)] \times [1 - \phi(k)]^{n-1}. \quad (6)$$

Equation 6 is the adaptation equation which will be used in applications.

### *Probability of a Response*

The probability of response  $A_1$  is the proportion of unadapted cues conditioned to it. The probability that cue  $k$  is unadapted is  $1 - a(k, n)$ , and the probability that it is conditioned to  $A_1$  is  $F(k, n)$ . Thus,

$$p(n) = \frac{\sum F(k, n)[1 - a(k, n)]}{\sum [1 - a(k, n)]}, \quad (k \text{ a member of } K). \quad (7)$$



### Validity of Cues

The experimenter determines the validity of a cue by differential reinforcement. Suppose that one response to cue  $k$  is reinforced proportion  $\pi$  of the time and the other is reinforced  $1-\pi$  of the time. The cue is a good predictor of reward if  $\pi$  is near one or zero, and it is useless if  $\pi = .5$ . The term "validity" underscores an analogy between the problem facing an  $S$  in selective learning and that facing a diagnostician. As the diagnostician tries to predict some criterion, using various test items with varying validities, so the  $S$  tries to select the correct response, using various cues which may be related in varying degrees to reward.

Validity of a cue,  $V$ , may be measured in several ways, but to get the simplest empirical laws it seems best to define it as

$$V = 4\pi^2 - 4\pi + 1. \quad (8)$$

The use of a quadratic function is suggested by the possibility that validity depends on the proportion of pairs of trials on which consistent information is received,  $\pi^2 + (1-\pi)^2$ , and on which contradictory information is received,  $2\pi(1-\pi)$ . The coefficients of the quadratic equation are fixed by making the validity of a consistently reinforced cue ( $\pi = 1$  or  $\pi = 0$ ) equal to one, and the validity of an irrelevant cue ( $\pi = .5$ ) equal to zero.

### Rates of Conditioning and Adaptation as Functions of Cue-Validity

In the writer's discrimination learning theory (11) it was assumed that the rates of conditioning of relevant cues,  $\theta$ , and adapting irrelevant cues,  $\phi(k')$ , were equal to one another and were also equal to the *proportion of relevant cues*. This basic idea is now generalized to experiments with vari-

ous reinforcement probabilities and to cues with validities between zero and one.

It is assumed that the rate-of-learning parameter controlling conditioning of cues,  $\theta$ , is equal to the mean validity of cues. That is,

$$\theta = \Sigma V_k / N, \quad (k \text{ a member of } K). \quad (9)$$

If all cues have validities 1 or 0, as is the case in the discrimination theory, then the mean validity is equal to the proportion of relevant cues.

It is assumed that the probability of adapting cue  $k'$  depends on the number of cues more valid than  $k'$ , and the degree to which they are more valid. Mathematically

$$\phi(k') = \Sigma (V_k - V_{k'}) / N, \quad (k, k' \text{ in } K, \text{ and } V_k \geq V_{k'}). \quad (10)$$

If  $V_{k'} = 0$ , then  $\phi(k') = \theta$  and the probability of adapting an invalid cue is the same as the probability of conditioning. If there is no cue with higher validity than  $k'$ , then  $\phi(k') = 0$  and  $k'$  is not adapted.

### RELATION TO STATISTICAL LEARNING THEORY

It was stated that the present theory is, under certain circumstances, very like Estes' statistical learning theory. If all relevant cues have the same probability,  $F(k, n)$ , of being conditioned to the  $A_1$  response, and if all irrelevant cues are adapted, then from Equation 9

$$\phi(n) = F(k, n). \quad (11)$$

If we substitute  $\phi(n)$  for  $F(k, n)$  in the difference Equations 2 and 3, we have the simplest cases of statistical learning theory. From Equation 7 we see that as  $n$  increases,  $a(k', n)$  asymptotically approaches 1, so that after sufficient training all irrelevant cues will be

adapted. Thus as training progresses the present theory gets closer and closer to the Estes model, and the asymptotes of the Estes model are in at least some cases the asymptotes of the present theory. In analysis of the present cue-learning theory it is often convenient, if only asymptotic probabilities are of interest, to study Estes' simpler model.<sup>4</sup>

This correspondence between the present theory and statistical learning theory is neither an accident nor a convergence of independent theoretical approaches. The elegance and empirical success of the Estes and Bush-Mosteller formulations led the present writer to adopt their fundamental notions and formulas, and the correspondence to earlier models was built into this theory.

#### RELATION TO DISCRIMINATION LEARNING THEORY

In an earlier paper (11) the writer offered a theory of two-choice discrimination with consistent reinforcement of relevant cues. A discrimination learning problem was described as having  $r$  relevant cues (with validity  $V_k = 1$ ) and  $i$  irrelevant cues with validity  $V_{k'} = 0$ . At Trial 1 neither relevant nor irrelevant cues are differentially conditioned—a situation we may represent in the present notation as

$$F(k,1) = F(k',1) = .5. \quad (12)$$

$$p(n_0+n) = \frac{r\{\pi - [\pi - F(k,n_0)](1-\theta)^{n-1}\} + iF(k',n_0+n)[1-a(k',n_0)](1-\theta)^{n-1}}{r+i[1-a(k',n_0)](1-\theta)^{n-1}}. \quad (17)$$

<sup>4</sup>An important advantage of Estes' and Bush and Mosteller's theories is their mathematical development, which gives detailed statistics of data in a wide variety of random reinforcement situations. The equations for  $p(n)$  of the present theory are considerably more difficult to work with in most cases.

Similarly, no cues are adapted at Trial 1, so that

$$a(k,1) = a(k',1) = 0. \quad (13)$$

From Equation 9,

$$\theta = r/(r+i) \quad (14)$$

and from Equation 10,

$$\phi(k) = 0 \quad (15)$$

$$\phi(k') = \theta. \quad (16)$$

These bounding conditions inserted in Equations 4, 6, and 7 yield the equations of discrimination learning theory.

#### VALIDITY AND THE RATE OF LEARNING IN GUESSING EXPERIMENTS

Several experiments on guessing as a function of the probability of reinforcement have found that the rate-of-learning parameters depend on  $\pi$ . Our notion of cue-validity, and the assumption that learning rates depend on validity (Equations 9 and 10), are intended to give a quantitative description of such findings.

In each such experiment we assume that there are  $r$  relevant cues with reinforcement probability  $\pi$ , and  $i$  irrelevant cues with reinforcement probability .5. Let  $k$  be a typical relevant cue, and  $k'$  a typical irrelevant cue. Substituting Equations 4 and 6 into Equation 7 and noting that  $\phi(k') = \theta$ , we have

If we let  $\delta = r/(r+i)$  be the proportion of relevant cues, it follows from Equation 11 that

$$\theta = \delta V \quad (18)$$

where  $V$  is defined by Equation 8.

If we divide all the terms in the right member of equation 17 by  $r + i$  and use the definition of  $\delta$  to simplify, and at the same time substitute  $\delta V$  for  $\theta$ , it follows that

$$p(n_0 + n) = \frac{\delta\{\pi - [\pi - F(k, n_0)](1 - \delta V)^{n-1}\} + (1 - \delta)F(k', n_0 + n)[1 - a(k', n_0)](1 - \delta V)^{n-1}}{\delta + (1 - \delta)[1 - a(k', n_0)](1 - \delta V)^{n-1}}. \quad (19)$$

Inspection of Equation 19 shows that if  $(1 - \delta V)^{n-1}$  approaches zero the limit of  $p(n_0 + n)$  is  $\pi$ . In such a situation the subject approaches the "probability matching" asymptote, eventually making response  $A_1$  as frequently as it is reinforced. The larger  $V$  is, the more rapidly will this asymptote be approached, for the larger  $V$  the smaller  $(1 - \delta V)^{n-1}$ . Thus Equation 19 relates the speed of learning to validity, and through validity to probability of reinforcement, in an appropriate fashion.

The aim of this paper is to make accurate quantitative statements about the relationship of learning rate to  $V$ . To do this we must establish the starting conditions by determining for some trial,  $n_0$ , the values of  $F(k, n_0)$ ,  $F(k', n_0)$ , and  $a(k', n_0)$ .

#### *Starting Conditions in Guessing Experiments*

In the experiments that we wish to consider, subjects commonly show a noticeable bias, so that  $p(1) \neq .5$ . It is possible to estimate  $p(1)$  from the proportion of  $A_1$  responses made on early trials. We do not know which cues are responsible for any bias that appears, so for convenience we assume that the initial bias is distributed equally over all cues. That is, we assume as a starting condition that

$$F(k, 1) = F(k', 1) = p(1). \quad (20)$$

Besides initial biases, one further consideration arises in regard to start-

ing conditions. While the experimenter may intend to reinforce alternative  $A_1$  on some fixed proportion  $\pi$  of the trials, he can actually accomplish this only on the average. On

the first trial he must reinforce one response or the other, and the actual relative frequency of reinforcements of  $A_1$  becomes either 1 or 0. It may happen that the subject gets only reinforcements of  $A_1$  and no reinforcements of  $A_2$  for the first four or six or even ten trials, though some  $A_2$  reinforcements are to be given. During this initial block of trials the problem given the subject is just the same as if the experimenter intended to give reinforcement consistently with  $\pi = 1$ . Thus, for some initial block of  $n_0$  trials, the subject works on a problem with actual  $\pi = 1$  and  $V = 1$ , and then shifts over to a problem with intermediate reinforcement probabilities.

An exact learning curve  $p(n)$  can be written using the actual sequences of reinforcements. Computations are easier if we assume that, for the first  $n_0 = \pi/(1 - \pi)$  trials, only  $A_1$  reinforcements are given ( $\pi > .5$ ), on Trial  $n_0 + 1$  an  $A_2$  reinforcement is given, and on all subsequent trials  $\pi$  is a satisfactory approximation to the actual relative frequency of  $A_1$  reinforcements.

#### *Probability of an $A_1$ Response on Trial $n$ in a Guessing Problem*

During the first  $n_0$  trials the subject has a problem just like simple discrimination learning with starting probability  $p(1)$ . Thus for any  $n$  less than or equal to  $n_0 + 1$ ,



$$p(n') = \frac{\delta\{1 - [1 - p(1)](1 - \delta)^{n'-1}\} + (1 - \delta)p(1)(1 - \delta)^{n'-1}}{\delta + (1 - \delta)^{n'}}, \quad (21)$$

whereas for any later trials  $n_0 + n$  we employ Equation 19 with

$$F(k, n_0 + 2) = \{1 - [1 - p(1)] \times (1 - \delta)^{M_0}\} (1 - \delta)$$

$$F(k', n_0 + 2) = .5 - (.5 - p(1)) \times (1 - \delta)^{n_0+1}$$

$$a(k', n_0 + 2) = 1 - (1 - \delta)^{n_0+1}.$$

In this fashion we construct a theoretical learning curve for learning in a guessing experiment.

In order to estimate parameters and test the theory conveniently, we wish to avoid the laborious process of fitting Equations 19 and 21. To do this we use the theoretical "total score" (as in 11, Equation 8), obtained by cumulating the learning function  $p(n)$ . The method, accurate to a good approximation, is to replace the discrete function  $p(n)$  by a function of a continuous variable  $t$ , and integrate the resulting  $p(t)$ . The integration is elementary<sup>5</sup> and gives an explicit function which, while bulky, is usable. We shall call this integral (which is not given here)  $T_m$  representing the cumulative  $p(n)$  from 1 to  $m$ .  $T_m$  is the total number of  $A_1$  responses made on the first  $m$  trials of training.

#### *Empirical Estimates of V as a Function of $\pi$*

If two or more groups of subjects from the same population are trained with the same apparatus and instructions but with different reinforcement probabilities, we assume they have the same cues available, and thus the same proportion of relevant cues  $\delta$ , though the validity of the relevant cues will differ. In such an experi-

ment one group of subjects can be used to estimate  $\delta$ , and this same  $\delta$  can be used for the other group, permitting a direct numerical estimate of  $V$  as a function of  $\pi$ . Grant, Hake, and Hornseth (9) ran five groups of twenty subjects each in a guessing experiment with reinforcement probabilities of 1.00, 0.75, 0.50, 0.25, and 0.00. Symmetry in the data permits us to lump the 1.00 and the 0.00 groups, calling them  $\pi = 1$ , and the .75 and .25 groups, calling them  $\pi = .75$ . The .50 group showed virtually no learning, as expected, and is disregarded.

The data from this experiment are used to check the assumed validity function given in Equation 8. The relevant cues for the  $\pi = 1$  group were assigned validity  $V = 1$ . From this group we estimated the proportion of relevant cues,  $\delta$ . Using this value  $\delta$ , we then found the value of  $V$  which would cause the .75 group to make 34.5  $A_1$  responses in 60 trials, as observed. The results of the computations are shown in Table 1. The obtained value of  $V$  at .281 is relatively close to the value predicted by Equation 8, which is  $4(.75^2) - 4(.75) + 1 = .250$ . There seems to be no way to test the statistical reliability of the discrepancy.

A second measure of  $V$  is accomplished by using the data from an experiment by Goodnow and Postman (8). They used reinforcement probabilities of 1.0, .9, .8, .7, .6, and .5 in a "problem-solving" situation. A peculiarity of their technique was a pretraining session in which the subjects were instructed and trained to attend to the relevant cues. We shall assume that this pretraining had the same effect as adapting all irrelevant

<sup>5</sup> See Burington's Tables (1, p. 82), formulas 309 and 310.

TABLE 1

SUMMARY OF CALCULATIONS ON GRANT-HAKE-HORNSETH EXPERIMENT

$\pi$	$p(1)$	$n_0 = \pi/(1-\pi)$	$\delta$	$T_{10}$	Estimated $V$	Predicted $V$
1.00	.46	—	.137	52.7	—(Assumed to be 1.0)—	
.75	.46	3	.137	34.5	.281	.250

cues before the actual training trials began. The assumption is stated as

$$a(k', 1) = 1$$

for any irrelevant cue  $k'$ . With this assumption we see that, according to Equation 7, for all trials  $n$ ,

$$p(n) = F(k, n).$$

With this special assumption it is possible to sum  $p(n)$  directly, dispensing with the approximation.

In analysis of the Goodnow-Postman experiment we again use the  $\pi = 1$  group to estimate  $\delta$ , the proportion of relevant cues. For all other groups all parameters except  $p(1)$  were known, and estimates of  $p(1)$  were made easily, since all groups showed about the same bias. Under these circumstances it is possible to make empirical estimates of  $V$  as a function of reinforcement probability. This was done using the .9, .8, .7, and

.6 groups. The detailed results of the computations are shown in Table 2. The correspondence between estimated and predicted values of  $V$  is quite good, especially for those groups which learned slowly and had small values of  $V$ . Since there were only eight subjects per group the estimates of  $V$  are probably not highly accurate. However, when the estimates of  $V$  from the Grant-Hake-Hornseth and Goodnow-Postman experiments are plotted against the prediction from Equation 8, as in Figure 1, the results are encouraging.

If we accept Equation 8 as representing the relationship between validity (and thus learning rate) and  $\pi$ , the theory can be used to explain the detailed results of guessing experiments reported by Estes and Straughan (7) and Jarvik (10). Estes and Straughan trained 16 subjects in each of two groups with reinforcement probabili-

TABLE 2

SUMMARY OF CALCULATIONS ON GOODNOW-POSTMAN EXPERIMENT

$\pi$	$p(1)$	$n_0 = \pi/(1-\pi)$	$\delta$	$T_{10}$	Estimated $V$	Predicted $V$
1.0	.43	—	.078	73.7	—(Assumed to be 1.00)—	
.9	.43	9	.078	66.7	.860	.640
.8	.43	4	.078	54.5	.260	.360
.7	.43	2	.078	48.0	.161	.160
.6	.43	2	.078	42.0	.041	.040

TABLE 3  
SUMMARY OF CALCULATIONS ON THE ESTES-STRAUGHAN  
AND JARVIK EXPERIMENTS

Estes-Straughan Experiment						
$\pi$	$p(1)$	$n_0$	$\delta$	Theoretical $V$	Total $A_1$ Responses (Observed)	Total $A_1$ Responses (Theoretical)
.85	.50	6	.167	.490	95.8	—
(.70)	.42	2	.167	.160	67.3	67.8
Jarvik Experiment						
$\pi$	$p(1)$	$n_0$	$\delta$	Theoretical $V$	Total $A_1$ Responses (Observed)	Total $A_1$ Responses (Theoretical)
.744	.46	3	.342	.238	60.0	—
.663	.46	2	.342	.106	52.0	49.8
.593	.46	2	.342	.035	46.5	46.8

ties of .85 and .30. Our computation is to estimate  $\delta$ , the proportion of relevant cues, from the data of the .85 group, and then assuming that  $\delta$  is the same for the .30 group and  $V$  depends on  $\pi$  according to Equation 8, compute the theoretical performance of the .30 group. Jarvik had three groups with reinforcement probabilities of .744, .663, and .593. The data of the .744 group are used to estimate  $\delta$ , and then performance curves are predicted for the other two groups according to the theory. The detailed results of these computations are shown in Table 3. Inspection of the results shows that the theory is able to predict the total  $A_1$  responses in  $m$  trials with acceptable accuracy in these two experiments.

Figure 2 summarizes the empirical accuracy of the proposed theory. The group with the highest reinforcement probability in each experiment is used to predict the results of the other groups, and theoretical expectations

are compared with observed data. In each of the predictions the starting probability  $p(1)$  was estimated from the data of the group predicted. Other than this, the predictions are computed without reference to the performance of the group predicted.

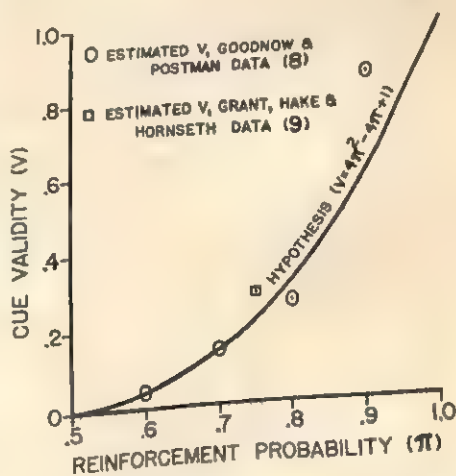


FIG. 1. Cue validity as a function of reinforcement probability—theoretical hypothesis (Eq. 8) and empirical estimates.



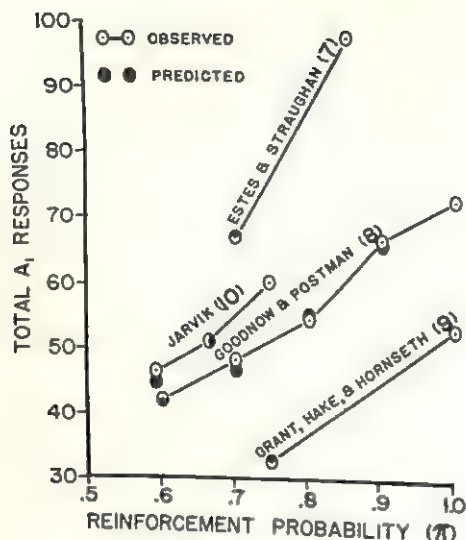


FIG. 2. Theoretical predictions of total-response scores in four experiments with empirical observed means.

### DISCUSSION

The theoretical formulation presented in this paper has been shown to give rise to the same asymptotes as does Estes' theory in simple probability-learning situations, to have the same consequences as the writer's earlier discrimination learning theory when applied to such data, and to give a quantitative account of the relation between reinforcement probability and speed of learning in guessing experiments.

As this theory describes learning, the probability of learning a particular cue depends on what other cues are present and on the correlations of those cues with reinforcements. This assumption is quite different from the approach taken by Estes and Burke (6), who suppose that each element has its own invariant sampling parameter. If the Estes-Burke model is right, a complex problem can be predicted theoretically as an assembly of stimulus elements, each with its own fixed parameter. Using the present

theory, we consider a complex problem as an assembly of elements which may appear in simpler problems, but the parameters of these elements depend on the context of the problem as a whole. While analysis of a problem as a whole seems at first too difficult, the results reported in this paper indicate that appropriate simplifying assumptions may make definite computations feasible.

The empirical results show that the present theory correctly describes the relation of learning rate to reinforcement probability. These results do not constitute evidence that the theory is correct, since the theory was formulated in the light of those results. In particular, the function relating validity to reinforcement probability, Equation 8, was selected because it seemed to fit the data. It is hoped that this formulation will stimulate the collection of more precise parametric data regarding the rate of learning under conditions of probable reinforcement.

### REFERENCES

1. BURINGTON, R. S. *Handbook of mathematical tables and formulas*. (3rd ed.) Sandusky, Ohio: Handbook Publishers, 1955.
2. BUSH, R. R., & MOSTELLER, F. A model for stimulus generalization and conditioning. *Psychol. Rev.*, 1951, 58, 413-423.
3. BUSH, R. R., & MOSTELLER, F. *Stochastic models for learning*. New York: Wiley, 1955.
4. ESTES, W. K. Toward a statistical theory of learning. *Psychol. Rev.*, 1950, 57, 94-107.
5. ESTES, W. K. Theory of learning with constant, variable, or contingent probabilities of reinforcement. In press, *Psychometrika*.
6. ESTES, W. K., & BURKE, C. J. A theory of stimulus variability in learning. *Psychol. Rev.*, 1953, 60, 276-286.
7. ESTES, W. K., & STRAUGHAN, J. H. Analysis of a verbal conditioning situation in terms of statistical learning theory. *J. exp. Psychol.*, 1954, 47, 225-234.

8. GOODNOW, J. J., & POSTMAN, L. Probability learning in a problem-solving situation. *J. exp. Psychol.*, 1955, 49, 16-22.
9. GRANT, D. A., HAKE, H. W., & HORNSETH, J. F. Acquisition and extinction of a verbal conditioned response with differing percentages of reinforcement. *J. exp. Psychol.*, 1951, 42, 1-5.
10. JARVIK, M. E. Probability learning and a negative recency effect in the serial anticipation of alternative symbols. *J. exp. Psychol.*, 1951, 41, 291-297.
11. RESTLE, F. A theory of discrimination learning. *Psychol. Rev.*, 1955, 62, 11-19.

(Received August 23, 1956)

# VALUE AND THE PERCEPTUAL JUDGMENT OF MAGNITUDE <sup>1</sup>

H. TAJFEL

*University of Durham, England* <sup>2</sup>

Problems connected with the influences exercised by motivational factors on perception have been subject to much debate in recent years. It seems that the dust of controversy is beginning to settle. A reaction is setting in (e.g., 17) against premature attempts to construct theories which would cope with the enormous variety of facts in this field, even before many of the facts have been properly established. These general theories were based in part on an *a priori* assumption that all the phenomena concerned can be reduced to, and formulated in terms of, a few very wide and nonspecific principles. No one can tell at present whether such an assumption is justified; but its hasty acceptance seems to have led to a confusion of experimental issues which has not been avoided either by those who stressed the importance of motivational variables in perception or by those who persistently denied it. The primary concern of this paper is with only one of these issues: the phenomena of perceptual overestimation, presumably due to value or some related variables. One of the reasons for this concern is that there exists a body of positive experimental findings in this field which cannot be ignored, and for which a consistent explanation has not yet been offered.

In the last fifteen years or so about twenty experiments on various aspects of overestimation have been conducted. Of

those, only two have yielded unambiguously negative results (3, 21). Partly negative results have been reported by Carter and Schooler (8), and by Klein, Schlesinger, and Meister (19). All other workers were able to conclude that, in the situations which they were using, "motivational" or "value" variables had an effect on their subjects' perceptual judgments of magnitude. Shifts in judgments of size (e.g., 1, 2, 5, 6, 7, 10, 20, 26), weight (e.g., 9), number (e.g., 25), and brightness (13) have been reported.

All this evidence cannot be dismissed as an accumulation of experimental artifacts. Furthermore, perceptual accentuation need not be considered as a "maladaptive" phenomenon. The fact that it may represent a departure from the "objective reality" leads to a general criticism which consists in pointing out that in order to survive we must perceive the world as it is, that we usually do, and that therefore the fleeting phenomena of overestimation are more typical of the specific laboratory situations in which they have been demonstrated than of perception under normal conditions. There is, however, a possibility that the shifts in the judgments of magnitude which concern us here do not interfere with an adequate handling of the environment. They may even be of help.

## A CLASSIFICATION

The experiments on overestimation fall naturally into two classes. In one group, changes in the magnitude of the stimuli under investigation are relevant to the changes in value. The experiments on coins provide an example here: in general, the larger the coin, the greater

<sup>1</sup> I wish to express my gratitude to Professor J. S. Bruner for encouragement and advice; and to Dr. W. Sluckin and Mr. D. Graham from the University of Durham for many useful suggestions.

<sup>2</sup> Now at the University of Oxford, Delegacy for Social Training.



its value. The experiment of Dukes and Bevan (9), in which judgments of weight of jars filled with sand were compared with judgments of weight of jars filled with candy, would also be "relevant" in this sense: heavier jars would presumably contain more candy, and thus have greater "value." On the other hand, several experiments have been reported in which changes in value have in no apparent way been related to changes in the physical dimension which the subjects were requested to judge. Thus, in the experiment by Lambert, Solomon, and Watson (20), the color of the disc was the determinant of its value, as red discs only were associated with reward; but judgments of size were requested of the subjects. In the experiments by Bruner and Postman (6) and by Klein, Schlesinger, and Meister (19), discs containing a swastika were among the stimuli used. Judgments of their size were compared with judgments of size of discs containing neutral symbols. Here again, the size of the swastika has no easily conceivable relationship to its degree of relevance. Further examples from both classes of experiments could be given.

#### "RELEVANT" DIMENSIONS AND THE ACCENTUATION OF DIFFERENCES

Let us first deal with the "relevant" group of experiments. The stimuli whose magnitudes were judged in these investigations form, by definition, a series varying concurrently in at least two dimensions: the physical dimension (size or weight), and the "dimension" of value. The concern of most experimenters has been to show (or to deny) that the stimuli of the series in which a variation in magnitude was paired in the environment with a variation in value were judged larger than stimuli of objectively equivalent magnitudes belonging to a different, "neutral," series. Thus, judgments of size of coins were

compared with judgments of size of cardboard or metal discs, and judgment of weight of jars filled with candy with similar judgments of jars filled with sand. Little attention has been paid to the perceived differences *between* the magnitudes of the stimuli in the "valued" series, as compared with the corresponding relationships in the "neutral" series.

It seems that some, at least, of the apparent contradictions between the various experimental results can be resolved if this intraserial aspect of the situation is considered. On the basis of an argument to be developed later, a prediction can be made that in a "relevant" series, where value changes concurrently with the dimension subjected to investigation, the differences between the stimuli of the series will be perceived as larger than the objectively equivalent differences between the stimuli of a neutral series, where no such association exists between value and magnitude. This comparison of the perceived differences within the two series would acquire its fullest meaning when intraindividual data are considered. Large individual variations are to be expected in the fineness of discrimination for both relevant and neutral series, but the tendency for a larger accentuation of differences in the former is likely to appear in most subjects, independently of the absolute values of the differences.

All the results to date in the "relevant" group of experiments support this contention. The quantitative difference, as expressed by subjects' judgments, between the two extremes of a series has been adopted as a rough measure of the accentuation of differences. This limitation was found necessary for two reasons. In the first place, coins, for example, differ not only in size and value. Secondly, it is possible that other determinants of quantitative judgment, such as the interaction of various gen-

TABLE 1

(CONVERTED FROM CARTER AND SCHOOLER, P. 202)\*

Stimuli	Percentage Increase from Actual to Perceived Extension
Cardboard discs	16.81
Aluminum discs	16.20
Coins judged when present	33.91
Coins judged from memory	38.69

\* Judgments of the "dime" and of the corresponding discs in other series have not been taken into account in this table, as the "dime" represents an exception in the concomitant variation of size and value in coins.

eralization gradients (cf., e.g., 18), might counteract in some segments of the series the expected accentuation of differences.

A glance at Fig. 1 in the early study by Bruner and Goodman (5) shows that the perceived difference between the smallest and the largest coins (extension of the scale of judgments) is much larger than the corresponding extensions for discs. A conversion of the data in the experiment by Carter and Schooler (8, Table I, p. 202) to similar extensions, for all subjects combined, shows that:

a. In all series of stimuli (coins, aluminum and cardboard discs) the perceived extensions are larger than the actual ones;

b. The relative differences between the actual and the perceived extensions are about twice as large for the series of coins as for both series of discs (see Table 1).

Carter and Schooler imply this when they "... suggest that there is a constant error involved in making these size judgments such that small coins are underestimated and large coins are overestimated in size" (8, p. 205); but they do not seem to draw further conclusions.

Bruner and Rodrigues (7) have introduced in their well-controlled experi-

TABLE 2

(CONVERTED FROM BRUNER AND RODRIGUES, TABLE 6, P. 21)

Stimuli	Percentage Increase from Actual to Perceived Extension	
	Value Set	Accuracy Set
Coins (on table)	68.2	73.7
Metal discs	34.9	55.7
Paper discs	20.9	39.8

ment the notion of "relative increase in overestimation" in which some of the suggestions made here are already implicit. Half of their subjects were assigned to a "value set" in which the instructions to the subjects emphasized the purchasing power of money. The other half were assigned to an "accuracy set" in which the experimenters' concern with accurate judgments of size was stressed. A conversion of their data similar to the above conversion of Carter and Schooler's data gives similar results (see Table 2).

This table is suggestive in several ways:

a. The same general tendency as in Carter and Schooler's data emerges: the relative extensions for the value series are considerably larger than for both neutral series.

b. All extensions in the "accuracy set" are larger than the corresponding extensions in the "value set." This is not inconsistent with the general proposition put forward here that the relevance of a series of stimuli is related to an emphasis of the differences between them. It is reasonable to assume that an appeal for accuracy of judgment would lead to similar results in its stress on clear discrimination between the stimuli of the series.

c. The differences between the extensions of the value and the neutral series

are less marked in the "accuracy" than in the "value" set. This is also consistent with the position adopted in this paper. An appeal for accuracy would have little effect in a series associated with value, where the accentuation of differences between the stimuli exists already as a long-standing product of past experience (cf. McCurdy's discussion of the role of schemata in the perception of coins, 22). This same appeal would have comparatively more effect in the case of a neutral series.

One further example: Dukes and Bevan (9) found that for their "positive" series (jars filled with candy) the variability of responses was less marked than for the "neutral" series (jars filled with sand). This is related to smaller j.n.d.'s in the positive than in the neutral series. In other words, the differences between the stimuli of the positive series were more clearly and consistently perceived than the objectively equivalent differences in the neutral series, and the scale of judgments was more extended for the positive than for the neutral series.

This summary of evidence suggests that the results of the "relevant" experiments on overestimation cannot be solely the product of some simple and rather mysterious process of overestimation. As has already been pointed out, the interest of the experimenters was confined mainly to the comparisons of the perceptual judgments of stimuli in a value series with the judgments of physically equivalent stimuli in a neutral series. The subjects were invited to make comparisons between the valued stimuli and either a neutral standard or elements of a parallel neutral series (with the exception of the experiment by Dukes and Bevan, in which comparisons *between* the stimuli of the value series were also included). However, the implications of the fact that the

stimuli associated with value do form a series cannot be ignored; during the experiments the subjects were repeatedly exposed to the various elements of this series, and the "belongingness," in the case of coins, would be further enhanced by familiarity.

In view of the evidence concerning the effects that all elements of a series, past and present, have on the quantitative judgments of its individual members (e.g., 14, 18), it may reasonably be assumed that the judgments of magnitude given by the subjects in the "relevant" overestimation experiments were not only determined by the perceived relationship, at the time of judgment, between a stimulus of the value series and a standard. They must have been affected as well by the background of perceived relationships between this particular stimulus and all other stimuli of the same series. This assumption is further supported by the evidence that the effects exercised by a particular stimulus on judgments pertaining to a series of stimuli increase as a function of the extent to which this stimulus is perceived as forming part of the series which is being judged (4).

In other words, in the "relevant" experiments on overestimation two aspects of the situation must be taken into account: the *interserial* and the *intraserial*. The first consists of the perceived relationships of magnitude between any stimulus of the value series and the neutral stimuli; the second is concerned with the perception of relationships between the stimuli of a value series as compared with the corresponding relationships in an objectively identical neutral series.

The phenomenon of accentuation of differences between the stimuli in a value series can be isolated in its "pure" form when the following two requirements are satisfied: the subjects' judgments must be based on comparisons between the



various stimuli of this series, and not between these stimuli and some extraneous ones; and the value and neutral series must be objectively identical, value being the only experimental variable in which they differ. These requirements were satisfied in some experiments recently completed by the writer (27).<sup>3</sup>

An ordered series of ten weights was used, and the subjects were requested to judge their heaviness in terms of seven category numbers. Each subject underwent an equal number of sessions under the "value" and under the "neutral" conditions. An experimental session consisted of two parts: in the first part, the entire series was presented several times in random order, but no judgments were reported by the subject. In the second part, following the first after an interval of about three minutes, all the stimuli were presented again several times, and judgments of weight were requested at each presentation. In the value condition, a small paper bonus (gift certificate, exchangeable for a book) accompanied each presentation of one of the two heaviest (or two lightest) stimuli of the series during the first part of the session. No rewards were given in the first part of the neutral sessions. In this way, the effects of the two experimental conditions, introduced in the first part of the sessions, on judgments of weight in the second part could be assessed. A total of 60 adult subjects were used in the four experiments. They were told, as part of the instructions, that the purpose of the experiment was to investigate the effects of monotony on the speed of performing a simple task which consisted of discriminating between weights. The "small paper tokens," which they were to receive from time to time, were being introduced in order "to vary the degree of monotony."

A questionnaire, which was presented to the subjects after completion of all the experiments, revealed that none of them doubted the truth of these statements. The subsequent disclosure of the experimenter's bad faith caused considerable surprise.

Data consisted of differences in the extensions of the scales of judgment between the two conditions for each of the subjects. Results can be briefly summarized as follows: in the first two experiments, in which rewards were associated with the two heaviest and the two lightest stimuli of the series, respectively, extensions of the scales of judgments were significantly larger for the value than for the neutral condition. In a third experiment, in which rewards were associated indiscriminately with all the stimuli of the series, no such effect occurred. In a fourth experiment, the procedure was the same as in the first and second experiments, apart from the fact that the paper "bonuses," passed by the experimenter to the subjects at each presentation of either one of the two heaviest or one of the two lightest stimuli, were devoid of all value. Once again, no significant effect on the extension of the scale of judgments was observed.

All this evidence is, to say the least, strongly suggestive. The advantages of using accentuation of differences as an explanatory device for overestimation are threefold:

a. It accounts for some seemingly contradictory results, such as the underestimation of the small end of the value series reported in some studies (7, 8); it also accounts for Bruner and Rodrigues' (7) "relative overestimation."

b. It does not require an introduction of *deus ex machina* principles to account for the phenomenon of overestimation in the "relevant" class.

c. With some extensions, it can be ap-

<sup>3</sup> This investigation was supported by a grant from the Durham Colleges Research Fund.

plied to the "interserial" differences in the "relevant" experiments, and to most experiments on overestimation which fall outside the "relevant" class.

The first point above does not require further elaboration. The second can be supported by some lines of argument, independent of each other.

McCurdy (22) has recently pointed out that by exaggerating the differences between the various coins, memory commits a "good error." In his ingenious attempt to relate the results of the studies on coins to the concept of schemata, he limited his discussion to series of coins where strong schemata could have developed through long familiarity with the stimuli. However, this would not account for the results of those experiments in which accentuation of differences in the value series, as compared with the neutral series, was found to exist (9, 27), in spite of the fact that the conditions for the development of the schemata were, as far as one can judge, not notably different for the two series.

What is the nature of the experience with coins, or with any series of stimuli, where it is important to discriminate sharply between the elements of the series? Discriminative responses to stimuli in such series are not usually made in terms of precise quantitative labels attached to individual stimuli. They are made in terms of "larger than," or "smaller than," the neighboring elements of the series. Minimizing the differences entails a risk of confusion; accentuating them is an additional guarantee of a successful response. However, the normal routine of responding in terms of "larger" or "smaller" is upset in most experiments on overestimation. The usual technique for correct handling of coins, for example, which is based primarily on an awareness of the relevant differences between a particular

coin and other coins of the series, present or absent, is not quite adequate, as unusually precise individual quantitative labels are requested in the experiments. These absolute labels, whether obtained through matching or through some kind of verbal categorizing, may be expected to reveal an accentuation of differences between the stimuli, since they would reflect, and possibly exaggerate, what is otherwise implicit in the relative judgments of comparison pertaining to the series.

It is in this sense only that the results yielded by the studies on overestimation can be treated as experimental artifacts. In this context, it is not particularly important to find out whether the stimuli are really "seen" as larger or smaller. They are *reproduced* as such; to ask the subject to match a variable standard to a stimulus, or to assign to the stimulus a quantitative verbal label, is essentially asking him to reproduce its size. This reproduction, which involves an activity very different from stating vaguely that an object is sizably larger or smaller than something else, lends itself easily to a sharpening of the relevant distinctive feature of it, which is, in this instance, its difference in size from the next object in the series. The phenomenon is not unfamiliar: Gibson reported some time ago (12) that a sharpening of differences occurred in the early stages of aircraft recognition training, when his subjects were asked to draw the silhouettes of the various aircraft.

A prediction is possible here to the effect that shifts in the estimates of magnitudes would either not occur or promptly disappear if the training to discriminate between the elements of a series, along which discrimination is important for the subject, was directed towards the accuracy of individual quantitative labels rather than towards a clear distinction in relative terms be-



tween the stimuli. This is exactly what happened in an experiment by Smith, Parker, and Robinson (25) in which accuracy of report was a condition of obtaining the reward. Prizes were offered to the subjects who would report correctly the greatest number of *dots* forming clusters which were flashed successively on a screen. In a control group, prizes went to those subjects who reported correctly on the number of dots in the greatest number of *clusters*. No information about the accuracy of their estimates was available to the subjects during the experiment. The first group showed overestimation in the early stages of the experiment, but after a certain number of trials the performance of both groups converged.

An experiment is being designed at present to test a further implication of the present argument: that in any ordered series, training to discriminate in comparative terms between the stimuli will result in larger estimated differences between the elements of the series than training to label the stimuli quantitatively in familiar units of measurement.

These assumptions of a "functional" mechanism underlying the accentuation of differences may be supplemented by some evidence coming from a different quarter. It consists of the findings about the effects of multidimensionality on the acuity of discrimination along a series. Eriksen and Hake (11) have reported recently that when the method of absolute judgments is used, the number of discriminable steps for a series of stimuli is greater when the stimuli vary concurrently in two or more dimensions than when they vary in one dimension only. More specifically, they found that discriminability was considerably greater when the "stimuli varied in size and hue, size and brightness, hue and brightness, and size, hue, and brightness" (p. 159) than when they varied in only one of these dimensions. Eriksen and Hake

suggest that their results were due to a kind of summation, to "the ability of Ss to make fairly independent judgments of stimulus values along each of the component dimensions" (p. 158). They add, however:

We cannot assume that this is always the case for compound stimuli. For some stimuli, judgments made of values in the separate dimensions may be interrelated. That is, the evocation of a particular response tendency by a component of a compound stimulus may change the likelihood of evocation of other response tendencies by other components of the stimulus. We could expect this to occur when Ss have learned by long experience to associate the occurrence of certain values in one dimension with the occurrence of particular values in another.

This interrelation may account in part for such perceptual "errors" as the size-weight illusion. Ryan (23, quoted by Hochberg, 15) writes: "... *O* is not really reporting *weight* at all, but *specific gravity*, and the illusion is really not an illusion at all. *O* is simply unable to report on weight being influenced by the striking differences in density involved." In the field of size estimation, Holzman and Klein (16) have recently reported that heavy discs were judged larger by their subjects than light discs of identical size, and that black discs were judged smaller than grey discs of identical size and weight. They add wistfully that these results "... do not exclude the possibility that value and need may indeed be important variables in size judgments" (p. 40).

Two conclusions can be drawn from the above: the first, supported by Eriksen and Hake's results, is that under some conditions of judgment, and in a situation new to the subject, compounding of concurrently varying dimensions will result in a clearer perception of differences between the stimuli along a series. The second is related to the first, and already outlined in the above quotation from Eriksen and Hake: when



the association between dimensions has been a long-standing one, new training may not be capable of inducing changes in discriminability. In such cases, judgments along one of the dimensions have already been influenced, previous to the experiment, by concomitant changes in other dimensions. In other words, it is likely that the Eriksen and Hake situation represents the incipient stages of a process which, if given an opportunity to develop further, might lead to interrelations such as those that Ryan suspects to be responsible for the size-weight illusion, or that Holzman and Klein (16) found between the estimated size and the actual weight of their discs. Interrelations of this kind exist, of course, when there is a consistent pairing, in the environment, of changes in more than one dimension.

Such pairing exists by definition in series where magnitude and value vary concurrently. The difficulty is that value is not a "dimension" in the physical sense. It is, however, an important attribute of the stimuli in such series, if only because efficient discrimination between the stimuli in terms of differences in value is, in most cases, more important than discrimination in terms of the physical dimension. It is not reasonable to expect that subjects who, in an experiment, judge a series of coins in terms of size, judge it in terms of size alone. This would be a feat of abstract behavior difficult for adults, and even more so for children. As Vernon (28) points out in her discussion of the experiment by Bruner and Goodman (5): "It seems fairly certain that . . . there was an inability to isolate a single aspect from a global percept" (p. 187). Or, as Hochberg writes in another context (15): ". . . the subjects are responding in terms of a dimension which specifies the most about the stimulus series with the least number of categories, rather than in terms

of the dimension expected by the experimenter."

In the experiment on weights described earlier (27), value is the only "dimension of difference" between the two series. It would also be a supplementary and important difference between two series which differ not only with regard to value, but also with regard to some other dimensions. In the first case, the difference in value is the only contributing factor to a more pronounced accentuation of differences in the value than in the neutral series; in the second, it is one of the contributing factors. Once again, this does not necessarily mean that the subjects "see" the stimuli in the value series as being more different from each other, or larger, or smaller. They respond *as if* they perceived them as such, as this is as far as the phenomenology of it can go for the time being.

#### "IRRELEVANT" DIMENSIONS AND THE ACCENTUATION OF DIFFERENCES

The accentuation of differences *between* the stimuli of a value series may partly account for the results obtained in the "relevant" experiments on overestimation. However, it has no direct bearing on the problems raised by the results of the "irrelevant" group of experiments; nor does it help to understand why in the "relevant" group a rather consistent trend towards overestimation of the value series should be superimposed, as it were, on the sharpening of differences within it.

The preceding discussion is not, however, entirely irrelevant to these problems. As was pointed out earlier, in a "relevant" series the subjects' judgments are simultaneously determined by the perceived differences between the stimuli of the series, and between a particular stimulus of the series and an extraneous stimulus. In the "irrelevant"

experiments, where the magnitude under investigation seems in no way related to the presence or absence of value, the fact that one stimulus is "valued" and another "neutral" would form an additional distinctive feature between them, superimposed on the physical differences. Examples of such experiments have been given above (6, 19, 20). Beams' study (2) on overestimation of size of favorite food objects by children provides another example. The common feature of these experiments is that the subjects are requested to judge along a dimension which, in usual circumstances, does not help to discriminate between the objects in terms of their value. Children, who in the experiment by Lambert *et al.* (20) were asked to report on the size of the chips, had just experienced a situation in which the value of the chips was determined by their color. The refugees, who in the study by Klein *et al.* (19) were judging the size of discs containing the swastika, would be unlikely, in any other circumstances, to pay special attention to the size of the emblem.

In these experiments the valued and the neutral objects belong to sharply distinctive categories of experience, and size is certainly not the basis of classification. In such cases, to judge size exclusively on the basis of size would, once again, be possible only if the subjects were able to abstract one element, and not an important one, out of a complex compound of experience. A more likely assumption is that the qualitative "motivational" differences between the two kinds of stimuli function in the direction of accentuating the perceived differences between their "irrelevant" magnitudes.

This would lead to the prediction that, all else constant, the perceived differences of magnitude between valued (or otherwise relevant) and neutral stimuli would be larger than the corresponding

differences between two sets of neutral stimuli. It does not, however, allow for the prediction that the valued stimuli would be perceived as *larger* than the neutral ones.

Some interesting indications which may suggest a way out of this difficulty can be found in the experimental literature. In the group of studies on "irrelevant" magnitudes, three experiments seem typical of those which yielded negative (or partly negative) results. Klein *et al.* (19) reported that their subjects did not reliably overestimate the discs containing the swastika. Bevan and Bevan (3), who used a "quasi-representative design" and a heterogeneous assortment of objects with their two children, found that the ratio of estimated to real size was not higher for the liked than for the disliked or neutral objects. Finally, Lysak and Gilchrist (21) found that increasing complexity of design was a significant determinant of increasing overestimation; but dollar bills were not judged larger than the corresponding rectangles, nor was there any correlation between overestimation and value of various dollar bills. Experiments by Lambert *et al.* (20) and by Beams (2), already referred to, are representative of those in which positive results were reported.

The experiments of Klein *et al.* (or at least the parts of them involving the swastika) and of Bevan and Bevan (3) have one feature in common: the value of the objects used has nothing to do with their size. The data, as they are presented, do not allow one to draw conclusions about an accentuation of differences between the sizes of valued and neutral stimuli; but if there was such an accentuation, there are no grounds to assume on the basis of the present argument that it should work, with any degree of consistency, in the direction of overestimating the valued objects.

On the other hand, it seems reasonable



to say that the dictum "the bigger the better" expresses a profound truth to Beams' 10- to 12-year-old Ss when they are confronted with a piece of fruitcake, a sugar cookie, or a "marshmallow-covered chocolate cupcake" (2, p. 197). Lambert *et al.* worked with much younger children (3 to 5 years old); the number of chips of a particular color which they received was related to the amount of candy which was the final compensation for their efforts. It would not be surprising if this, in view of the age of the subjects, had a great deal to do with their judgments of size.

The negative results of Lysak and Gilchrist (21) lend further support to the thesis that the accentuation of differences will tend consistently towards an overestimation of valued objects *only* when there exists a valued series in which there is some discernible relationship between magnitude and value. Their adult subjects knew from long experience that the value of dollar bills does not vary as a function of their size. They may have accentuated the differences in size between the bills and the corresponding rectangles, but not necessarily or consistently by overestimating the bills more than the control stimuli.

A comparison of this study with other experiments performed on adult subjects helps to bring out more clearly its particular features which may account for the lack of positive results. In the experiment by Klein *et al.*, there is no tendency to overestimate the disc bearing the swastika; but in one of their two experimental situations, all results pertaining to the disc bearing the dollar sign are significant (19, Table 3, p. 103). This is further substantiated by Solley and Lee (26) who attempted "... to determine whether the differences in perceived size of discs with symbols drawn upon them, as reported by Bruner and Postman, could be better explained by their hypothesis of symbolic value than

by the more general Gestalt principle of closure" (p. 142). They report an overestimation of the discs containing a dollar sign (as compared with discs containing a figure of presumably equal degree of closure), but no such effect with the discs containing a swastika. The fact that a symbol of money drawn on a disc is strongly suggestive of the familiar series of coins, whose value increases with their size, may well be responsible for these results.

The negative outcome of the experiment by Lysak and Gilchrist must also be taken in conjunction with an almost identical experiment by Dukes and Bevan (10), who reported positive results. They used a number of rectangular cards, identical in size but differing in value, and engaged in "gambling" with their subjects. A number representing its value was stamped on each of the cards. The cards were drawn at random from a bag, and at each draw the subject won or lost an amount of money proportional to the positive or negative number printed on the card. As soon as each drawn card was returned to the bag, the subject was asked to match its size with one of a series of blank "reference" cards spread out in front of him. Dukes and Bevan's hypothesis was that "as the monetary value printed on the test card increases from zero, the subject, regardless of winning or losing, will tend to select larger reference cards as his estimates" (10, p. 47). This was confirmed: a significant correlation emerged between the ratios of estimated to actual sizes and the monetary value, positive or negative, of the test cards.

The situation was almost identical with one of the arrangements of Lysak and Gilchrist's experiment (21). If the fact of conflicting results were to be ascribed, here or in other cases, to relatively unimportant differences in procedure, attempts to generalize results from a great number of psychological



experiments would have to suffer a serious setback. There is, however, one essential difference between the two experiments: Lysak and Gilchrist's subjects *knew* that all the dollar bills, with which they were presented, were of equal size. This bit of information about the cards used was not available to Dukes and Bevan's subjects. What is more, Dukes and Bevan included in their instructions the following sentence: "Here is a bag containing small white cards, *varying slightly in size*" (10, p. 44, italics mine). It would be only natural for the subjects to assume that this variation followed some regular pattern, and that this pattern consisted of an increase in size corresponding to the increase in value. Here again, a continuous variation along the dimension of value has carried over to the judgments of the physical dimension, and this has been further helped by a hint from the experimenter.

But there is some evidence that the same could have happened even without the hint. In an experiment by Ashley, Harper, and Runyon (1), a "grayish metal slug" was presented several times to the subjects. At different presentations of the slug they were told that "it was made of either lead, silver, white gold, or platinum" (p. 568). Judgments of the size of the slug reflected its imaginary value: the size of the variable standard "that was called equal to the slug increased as the cost of the metal increased" (p. 572). As in the preceding example, a progressive increase in value is related to an imaginary increase in size, while the actual size is kept constant.

The evidence summarized so far lends strong support to the view that both the positive and the negative results of the studies on overestimation can be best understood in terms of a hypothesis of accentuation of differences. This accentuation would work along two axes:

between the stimuli of a series in which there is a concomitant variation of value and magnitude; and between the valued stimuli and the neutral ones. In the latter case, the accentuation would tend towards a relative overestimation of the valued stimuli only under certain conditions, which have been discussed above.

#### SOCIAL PERCEPTION

Some implications of the present discussion go beyond the problems raised by the phenomenon of overestimation. It may be said that, in a sense, "overestimation" as discussed in this paper is a special case and a convincing experimental paradigm of a more general aspect of social perception. Many social objects and events are sharply classified in terms of their value or relevance. When judgments concerning some quantifiable or ratable aspects of stimuli which fall into distinct categories are called for, differences in value or relevance cannot fail to influence the quantitative judgments in the direction of sharpening the objectively existing differences between the stimuli. A very similar conception has recently been formulated by Hochberg (15), and applied by him to the perception of in- and out-group individuals:

If a group of individuals is perceived as different from the non-group individuals, the perceived differences between those within the group and those outside the group will automatically be sharpened, and the differences perceived between the members of the group (i.e., intragroup differences), and between those outside the group will be lessened.

These judgmental effects of categorization are probably fairly general; it is likely, however, that they are particularly pronounced when judgments are made in dimensions in which scaling in magnitude is simultaneously a scaling in value. Thus, it may well be that an accentuation of differences in size will

hardly occur between two paintings, one liked and one indifferent or disliked. But when skin color, or height, or some facial traits of social "value" are concerned, there will be a marked sharpening of differences in the degree of these characteristics perceived as belonging to individuals who are assigned to different categories. Some evidence of this is provided by a recent study on "perceptual accentuation and the Negro stereotype" conducted by Secord, Bevan, and Katz (24). Their results suggest that a group of prejudiced Ss sharpened the differences in the degree of Negroid physiognomic traits possessed respectively by Negroes and whites, more than did a group of nonprejudiced Ss. It is likely that the same is happening in the case of more abstract social judgments which are implicitly quantitative, such as, for example, those concerning the relative frequency of crimes in various social groups, as perceived by people who have an axe to grind.

#### SUMMARY

The thesis put forward in this paper is that the apparently conflicting results of studies concerned with perceptual overestimation can be best understood if this phenomenon is considered as a special case of accentuation of perceived differences. This would work in two directions. First, in a series of stimuli in which there exists a concomitant variation of value and of some physical dimension, the perceived differences between the elements of the series would be larger than in a corresponding neutral series. Secondly, the perceived differences in magnitude between the stimuli which possess the attribute of value and those which do not would tend to be more accentuated than the corresponding differences between neutral stimuli. This would not, however, always result in overestimation. The sharpening of

differences resulting in overestimation of valued stimuli would occur only when, in the class of stimuli with the attribute of value, an increase in some physical magnitude is correlated with the increase in value. In other cases, there are no grounds for assuming that an accentuation of perceived differences between the two classes of stimuli would tend consistently towards an overestimation of the valued stimuli.

In this way, overestimation is seen as one instance of a more general phenomenon. The accentuation of differences between classes of stimuli occurs when these stimuli differ in some respects other than the dimension along which the subjects are reporting their judgments of quantity. The presence or absence of "value" or "relevance" is one such contrast. It is probable that this contrast is responsible not only for the results obtained in the field of overestimation. It may also lead to the sharpening of differences between the quantifiable or ratable aspects of social objects and events. This occurs when the dimension along which judgments are made is not the primary basis for assigning the various stimuli into their sharply distinctive categories.

#### REFERENCES

1. ASHLEY, W. R., HARPER, R. S., & RUNYON, D. L. The perceived size of coins in normal and hypnotically induced economic states. *Amer. J. Psychol.*, 1951, 64, 564-572.
2. BEAMS, H. L. Affectivity as a factor in the apparent size of pictured food objects. *J. exp. Psychol.*, 1954, 47, 197-200.
3. BEVAN, W., & BEVAN, D. C. Judged size and personal relevance: an exercise in quasi-representative design. *J. gen. Psychol.*, 1956, 54, 203-207.
4. BROWN, D. R. Stimulus-similarity and the anchoring of subjective scales. *Amer. J. Psychol.*, 1953, 66, 199-214.
5. BRUNER, J. S., & GOODMAN, C. C. Value and need as organizing factors in per-

- ception. *J. abnorm. soc. Psychol.*, 1947, 42, 33-44.
6. BRUNER, J. S., & POSTMAN, L. Symbolic value as an organizing factor in perception. *J. soc. Psychol.*, 1948, 27, 203-208.
  7. BRUNER, J. S., & RODRIGUES, J. S. Some determinants of apparent size. *J. abnorm. soc. Psychol.*, 1953, 48, 17-24.
  8. CARTER, L. F., & SCHOOLER, K. Value, need, and other factors in perception. *Psychol. Rev.*, 1949, 56, 200-207.
  9. DUKES, W. F., & BEVAN, W. Accentuation and response variability in the perception of personally relevant objects. *J. Pers.*, 1952, 20, 457-465.
  10. DUKES, W. F., & BEVAN, W. Size estimation and monetary value: a correlation. *J. Psychol.*, 1952, 34, 43-53.
  11. ERIKSEN, C. W., & HAKE, H. W. Multi-dimensional stimulus differences and accuracy of discrimination. *J. exp. Psychol.*, 1955, 50, 153-160.
  12. GIBSON, J. J. Social perception and the psychology of perceptual learning. In M. Sherif & M. O. Wilson (Eds.), *Group relations at the crossroads*. New York: Harper, 1953. Pp. 120-138.
  13. GILCHRIST, J. C., & NESBERG, L. S. Need and perceptual change in need-related objects. *J. exp. Psychol.*, 1952, 44, 369-376.
  14. HELSON, H. Adaptation-level as a basis for a quantitative theory of frames of reference. *Psychol. Rev.*, 1948, 55, 297-313.
  15. HOCHBERG, J. E. Psychophysics and stereotyping in social perception. In *Emerging problems in social psychology*, in press.
  16. HOLZMAN, P. S., & KLEIN, G. S. Inter-sensory and visual field forces in size estimation. *Percept. Mot. Skills*, 1956, 1, 37-42.
  17. JENKIN, N. Affective processes in perception. *Psychol. Bull.*, 1957, 54, 100-127.
  18. JOHNSON, D. M. *The psychology of thought and judgment*. New York: Harper, 1955.
  19. KLEIN, G. S., SCHLESINGER, H. J., & MEISTER, D. E. The effect of values on perception: an experimental critique. *Psychol. Rev.*, 1951, 58, 96-112.
  20. LAMBERT, W. W., SOLOMON, R. L., & WATSON, P. D. Reinforcement and extinction as factors in size estimation. *J. exp. Psychol.*, 1949, 39, 637-641.
  21. LYSAK, W., & GILCHRIST, J. C. Value, equivocality and goal availability. *J. Pers.*, 1955, 23, 500-501. (Abstract)
  22. MCCURDY, H. G. Coin perception studies and the concept of schemata. *Psychol. Rev.*, 1956, 63, 160-168.
  23. RYAN, T. A. Interrelations of the sensory systems in perception. *Psychol. Bull.*, 1940, 37, 659-698.
  24. SECORD, P. F., BEVAN, W., & KATZ, B. The Negro stereotype and perceptual accentuation. *J. abnorm. soc. Psychol.*, 1956, 53, 78-83.
  25. SMITH, K. R., PARKER, G. B., & ROBINSON, G. A. An exploratory investigation of autistic perception. *J. abnorm. soc. Psychol.*, 1951, 46, 324-326.
  26. SOLLEY, C. M., & LEE, R. Perceived size: closure versus symbolic value. *Amer. J. Psychol.*, 1955, 68, 142-144.
  27. TAJFEL, H. The role of value in the formation of a scale of judgments. *Bull. Brit. psychol. Soc.*, 1956, 29, 14. (Abstract)
  28. VERNON, M. D. The functions of schemata in perceiving. *Psychol. Rev.*, 1955, 62, 180-192.

(Received September 21, 1956)



## A MECHANICAL MODEL FOR HUMAN ATTENTION AND IMMEDIATE MEMORY<sup>1</sup>

D. E. BROADBENT

*Applied Psychology Research Unit, Cambridge, England*

Many people have a natural distaste for model building. A mechanical model is essentially a theory expressed in material parts rather than in abstract symbols such as words or mathematical expressions. Its logical standing is the same: that is, it stands or falls by the degree to which it fits the results of experiments on human or other animals. Yet many models in the past have been somewhat undistinguished in the closeness with which experiment has been considered in their design. It is difficult to avoid feeling that this is because a model is unduly laborious to build as compared with a verbal theory, so that the builder tends to become obsessed with the properties of his model rather than those of the organism. Consequently there is much to be said for building theories verbally, and especially for using the qualitative terms of information theory in the hypothetico-deductive fashion ably set out by Mackay (29). Such an approach has certain advantages over the other popular alternative of quantitative S-R terms; these advantages have been considered elsewhere (11).

For example, the writer holds that the human perceptual system has a limited capacity, that in consequence a selective operation is performed upon all inputs to the system, and that this operation takes the form of selecting all inputs having some characteristic in common.

<sup>1</sup> This work was supported by the British Medical Research Council, and the writer works under the general direction of Dr. N. H. Mackworth. He has discussed the topic with profit with many individuals; Dr. J. Brown should be especially mentioned.

Such an operation extracts little information from the signal and thus should be economical of nervous mechanism. Characteristics on which the selection can operate may be named "sensory channels." The particular selection made at any one time will depend partly on characteristics of the input itself (physical intensity, earliness in time, absence of recent inputs on that channel, position of the channel in the hierarchy of all channels) and partly on information in a more permanent store. The change from one selection to another will take a determinate time.

Incoming information may be held in a more temporary store at a stage previous to that of the selective operation. Such information will pass through the perceptual system on the next subsequent selection of the sensory channel of its arrival, if it is still in store; but the probability of the latter condition being fulfilled will decline with time spent in store. After passage through the perceptual system, information may be returned to the same temporary store, the selection of information for such return being determined by information in a more permanent store.

But although such a purely verbal theory may fit experimental results, it is difficult to communicate to others without putting them to the trouble of learning the necessary vocabulary. And if the theory is rephrased, still abstractly, it is open to misinterpretation; thus Deese (19), in a paper which makes a number of valuable contributions to the theory of prolonged work, has described the writer as postulating an inhibitory construct in human per-

formance. This misunderstanding probably arises from the fact that the theory outlined in the last two paragraphs is intended to apply both to conditioning and to human watch keeping; in neither case is it thought necessary to find an inhibitory construct of orthodox learning theory type. To say that conditioning and perception are related is not to apply a particular interpretation of the former to the latter.

Clearly, then, some sort of expository device is needed for an abstract theory using unfamiliar terms. And it is even difficult for the theorist to remember in abstract form the results of the many different experiments which a good theory should consider. A simple mechanical model has the virtue of avoiding these difficulties. It has other vices: it may have accidental properties which mislead research. Perhaps the best compromise is to state a theory in abstract terms, and also to give a model which can be described by the same verbal theory. Information concepts are applicable to any system, whatever its physical nature, and so may equally fit a model or a man. This is the approach of Deutsch (20, 21), and it has real advantages, independent of the value of his particular theory. The present paper is therefore intended to describe an extremely simple model of the human perceptual system. It may serve both as an easy introduction to the formal theory in information flow terms and also as a convenient mnemonic for the results of a number of experiments.

### THE BASIC MODEL

The necessary requirements are a Y-shaped tube (Fig. 1) mounted vertically, and a set of small balls. Each ball bears a number so that all are individually recognizable. The Y tube has a narrow stem which will just take

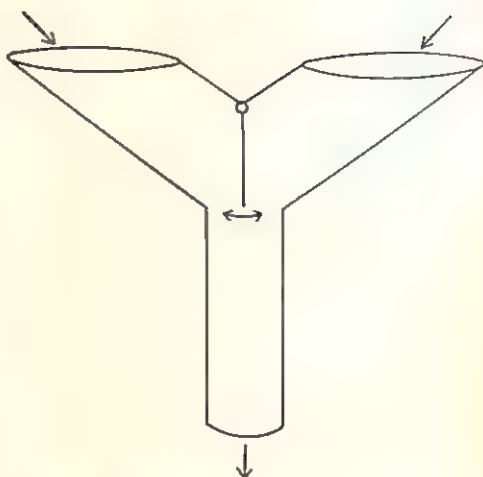


FIG. 1. The simple model for attention.

only one ball, though the branches are wider. At the junction of stem and branches is a hinged flap which normally hangs straight downward, but which can be pivoted about its upper edge so as to close off either of the branches of the Y. This pivoting can be done by a handle from outside the tube, controlled by stored information; purists may wish to control the handle by a punched-tape system, but a human being is an adequate substitute. When the handle is left alone the flap moves freely so that a ball dropped into one arm of the Y will knock the flap aside and fall into the stem of the Y.

In this model the balls represent the information from various stimuli. The branching arms represent different sensory channels; thus one might be one ear and one the other ear. Alternatively, one might be the ear and one the eye. (Sensory channel is not, however, quite equivalent to sense-organ, since we would treat sounds localized in different places as being on different channels.) The bottom of the Y represents a response output, so that the process of dropping a ball into the arms and observing its emergence at the bottom is analogous to that of delivering a stimu-

lus and observing a response. The behavior of the model resembles that of man in the following ways:

(a) If two balls are dropped simultaneously, one into each of the branches, they will strike the flap on both sides; it will not move and therefore they will jam in the junction. Numerous experiments show this "distraction" effect, but there are certain advantages in citing an auditory experiment (5, 33). These same experiments show that if the handle is used to shut off one branch before the balls are inserted, then the ball entering the other branch will emerge successfully, which is analogous to previous instructions.

(b) If the two balls are not strictly simultaneous, the first to arrive will obtain an advantage by knocking the flap over and shutting out the other. This had been shown to be analogous with competing auditory stimuli by Spieth, Curtis, and Webster (36).

(c) If the Y is not perfectly vertical, the ball in the more vertical branch will have an advantage over a simultaneous ball in the other because the door will hang to one side. Equally, one sensory channel may have an advantage over another, as has been shown for high-pitched noise as compared with low (10).

(d) If one ball is flung violently down its branch, it may succeed in forcing over the door against the unassisted weight of a ball on the opposite branch. Equally, an intense stimulus may have an advantage (2, 10).

(e) After a single ball has been passed through the system, the door will swing back from the position into which it has been pushed. Naturally it will overswing, and temporarily close the branch which has just been used. A stimulus has similarly an extra advantage for response if it comes on a previously quiet channel as opposed to a

previously busy one. This has been shown by Poulton (34) for auditory signals, and a related finding is that of Hyman (25) for visual reaction times. In the latter case, stimuli of different frequencies of occurrence were delivered, and it was found that average reaction time to a set of stimuli was proportional to the information conveyed by them; but the infrequent signals, while giving long reaction times, did not give times as long as those to be expected from information theory calculations. Note, however, that the time taken for a swing is important in the model; this will probably also be true in man, to judge from data which are best considered below. It may also be connected with Hyman's finding that, for numbers of alternatives greater than two, the second of two identical signals received an unduly fast reaction.

(f) If a given number of balls are to be put through the tube, it is best to deliver them asymmetrically, the majority to one branch. There is then less risk of jamming than if equal numbers are admitted by both branches. The analogous finding for auditory messages has been reported by Webster and Thompson (37). This point is related to another, that if balls are being inserted into one branch at random intervals the effect of increasing the rate of delivery of balls through that same branch (the "speed" of work) is not the same as that of adding the same number of extra balls to the other branch (the "load" of work). There is more risk of jamming in the latter case. This point was first clearly emphasized by Conrad, using visual signals (16). Obviously the effect of using two branches rather than one will be more serious if the rate of delivery of balls is high, since this increases the probability of a jam. Conrad showed such an interaction of speed and load. Mackworth and Mackworth (28) have shown a similar



effect in a different type of task, and have demonstrated that the fluctuating difficulty of this complex task at any instant can be represented by the amount by which each signal is overlapped by other simultaneous ones.

At this point we may pause to consider the cynical reader who is wicked enough to be doubtful of the existence of a Y-shaped tube somewhere in the region of the thalamus. Such a reader will probably have noticed already a situation which will follow when two balls are dropped into the model and jamming is avoided by the door being to one side for one of the reasons listed. The favored ball will descend the stem of the Y, but the impeded ball will not therefore disappear. It will emerge later, when the door next swings back to the opposite branch. Surprisingly enough, this also happens with man. Simultaneous stimuli either jam or produce successive responses.<sup>2</sup> It has been shown for three different combinations of sensory channels (7, 8), and is probably an effect identical with the "prior entry" of classical psychology. In the same way Conrad has not only shown complete failures to respond (17) but also shifts in the time at which responses appear (18).

Having thus demolished the cynic, we must add that a slight complication should be added to the model in order to cover perfectly the results of experiments on successive response to simultaneous stimuli. But this will be left

<sup>2</sup> At the risk of complicating the issue, we must say that "stimuli" in this sentence means "stimuli not of low information content." A familiar predictable sequence of stimuli may quite well produce responses simultaneous with other responses (1). And the same S-R unit interferes less with another task when drawn from a smaller ensemble of possibilities (12). This point is of course implied in the statement that the perceptual mechanism has a limited capacity in the sense of information theory.

until the section on immediate memory; let us first deal with the effects of speed and of prolonged performance on the simple model.

(g) As the stem of the Y is so narrow, there will be a certain amount of delay between the insertion and emergence of each ball. At slow rates of insertion each ball will, however, emerge before the next is inserted. At faster rates there will develop a lag such that one ball may be inserted before the previous one emerges. With still faster rates balls will begin to accumulate in the branch; response will get further and further behind stimulus as the task proceeds. But this is a desperate expedient and will lead to breakdown when the branch is full unless the rate slows down again. These stages have been shown by Vince (35) for visual stimuli; a related effect has also been shown in less detail in hearing (4). A result of the piling up of balls in the branch when the rate is too high is that a ball inserted at a very short interval after another will remain in the tube longer than is normal: the "psychological refractory period" (38). It must be remembered that balls are analogous to information and not to stimulation; highly probable stimuli will not give an unduly long reaction under these conditions (22).

(h) Now suppose that we wish to operate the handle in such a way that one branch always has priority: prolonged performance of a task involving only one sensory channel. If there is a fairly rapid flow of balls down the selected branch, the handle will hardly need holding after the initial setting, since the beginning of any swing back will probably be checked by another ball. But if the selected branch is not very busy, the weight of the door must be held by a positive force on the handle. In this case the fingers holding the handle will fatigue; for purists, the

punched-tape machine will only hold one branch shut for a limited period. When the handle is released the door will swing back to close the opposite branch, and then return pendulum-wise to the desired position when it will again be held. The result will be that prolonged tasks in which unexpected stimuli appear for less than a certain period (the "swings of the door") will show marked decrements in performance. But similar tasks in which the stimuli are present for longer times will show much less decrement.

The earlier evidence for this view of prolonged performance has been summarized elsewhere (6). But earlier presentations have misled Deese (19) into supposing that this view implied a decrement with continued performance of *any* vigilance task. He employed tasks in which a signal was either presented repeatedly until seen, or else painted on a tube face by a sweep line which did not return to obliterate the signal for nearly three seconds. It was in fact predicted by the writer (6, p. 300) that such tasks would show little or no decrement, and this was found by Deese to be the case. The latter type of task gave more signs of decrement, probably being a borderline case, since the trace left on the phosphor screen decayed until scarcely visible at the end of the three seconds. It should be noted that an expectancy theory of the type favored by Deese, while undoubtedly applicable to some aspects of vigilance, is not able to account for the effects of varying the length of signal presentation. In our Y-tube model, the role of expectancy is incorporated by describing the balls as representing information rather than stimulation; the more probable signals receive more efficient response. But the swinging of the flap is also necessary.

To summarize the writer's present views on vigilance, the efficiency of a

man asked to detect infrequent signals should be described by both a mean and a variance. The mean level is determined by such factors as the rate at which signals arrive but not necessarily by the length of time since the session began. The variance, on the other hand, increases as the session progresses, short intervals of extremely low efficiency being interspersed with fairly long periods of normal or supernormal efficiency. The score from any given task may depend on one or on the other of these quantities. Thus, for example, the earlier British work has been mostly concerned with the instants of very low efficiency; the type of results reported by Deese and by later British workers (6) mostly with the mean over appreciable periods of time.

To return to our model, two further points should be made. The first is that the door need not swing the whole way over when the handle is released; it may reach a central position just as two balls arrive. There will then be a jam rather than passage of the wrong ball. Equally, failure of reaction to a task need not imply overt reaction to some irrelevant stimulus.

Secondly, as the balls represent information rather than stimulation, tasks in which the sequence of stimuli is predictable will not show fatigue decrements of this type. To show decrements, the signals must be unpredictable either in content, as were those of Bills (3), whose "blocks" are instances of this effect, or else in time of occurrence, as were those of Mackworth (27); this, however, is only one of the necessary conditions.

#### IMMEDIATE MEMORY

Our Y tube does seem at this stage to have related a number of facts about perception and put them in a way which most people can understand. It is admittedly ludicrous as a description of

what really happens in the brain, but this is a positive advantage. Psychologists are not likely to mistake this model for speculative neurology, and so they should concentrate their experiments on the essentials of the theory rather than the irrelevant properties of the model. As a device for communicating the outline of the theory, however, the model seems sufficiently adequate to justify an extra complication in order to express a theory of immediate memory. This theory is in a slightly different position from the views on perception which have been given so far; the latter are entailed by the experiments, but the theory of immediate memory is not the only conceivable explanation of the observed facts. Yet it has a fairly high probability and is worth discussing.

The complication is twofold. First, a device must be supposed fitted to each branch of the Y tube, such that if any individual ball remains in the branch for more than a certain time continuously it is removed from the system completely. This could be done by filling the tube with acid, but the writer does not wish to encourage the development of a race of fingerless psychologists. Mechanical devices are quite possible, though complicated to describe, and the details will therefore be left unspecified and available from the author. The second complication is that from the foot of the stem two return tubes lead back to the branches (Fig. 2). Admission to these return tubes is controlled by a lower door which again is operated by a handle. The latter is dependent on an outside operator, or the familiar punched-tape machine—in brief, some form of stored information. Finally, as the return tubes are operating against gravity they must contain some form of conveyor, but this has no particular psychological significance.

In immediate memory experiments, as usually performed, a stream of stimuli

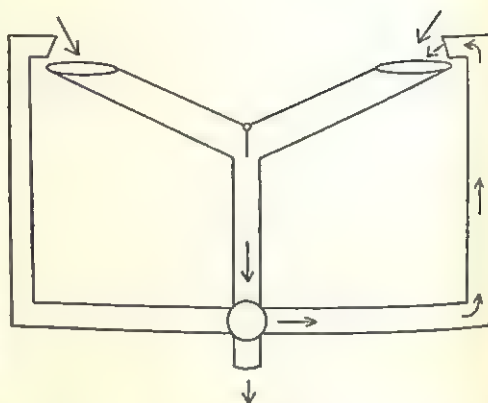


FIG. 2. The model modified to illustrate the theory of immediate memory as a recurrent circuit; or, in other terms, as a fading trace periodically revived by rehearsal.

is delivered completely before any overt response is required. Thus all the information in the stream is within the organism at one time; all the balls are somewhere in the tubes. Yet with small numbers of balls the time between insertion and emergence can be varied without any apparent effect on efficiency (13). Let us consider some of the methods which may be adopted.

(a) In the first place, let us suppose a short series of balls to be waiting in one of the branches. If they stay there indefinitely, the critical time will be exceeded and the ball which has been there longest will leave the system. Thus the branch will not serve as a store during an unlimited period of delay before response is allowed. What is perfectly possible, however, is for the series of balls to pass down the stem and, by the lower door, back up a return tube to the other branch. As soon as the full series has completed this round the process can be repeated, so that the set of balls can be kept in the system for any desired time. But note that insertion of any other ball during that time will either mean that the extra ball will never emerge, or else that the circulation of the series will meet inter-



ference. Note also that the system will only operate with a limited number of balls; above that number each ball will have to wait in the branch for more than the critical time, and there will therefore be a severe breakdown of storage. These are well-known characteristics of immediate memory. Less obvious is the suggestion that the interfering effect of an extra ball during the delay will depend on the size of the series circulating, being most serious near the limit. This has been shown by Brown (14). He also demonstrated that the interval before or after the interpolated stimuli had little effect, as one would expect, and that if the extra stimuli came before the memory span stimulus there was little effect. It is Brown's explanation of his results which has prompted the present account.

(b) If both branches contain balls, movements of the door will put one group into circulation before the other; to move the door over while a group is actually passing would be to risk a jam. So all balls on one branch will pass down the stem before any on the other. As has been said, the memory analogy is in fact true, if a digit memory span is obtained with half the digits on one ear and half on the other, or half on the eye and half the ear (7, 8). In addition, the first branch to be dealt with undergoes less risk of any ball reaching the critical time; equally, performance is better on the first set of digits to be recalled (9, 13).

(c) So far we have considered only the case in which the branches start with balls in them. But in practice these balls will be inserted into the branch one after another, and the extra time intervals thus produced will have their effect. Thus, for example, if balls are inserted rapidly into both branches simultaneously, the difficulty of moving the door will apply, as mentioned in the last paragraph; and so all the balls in

one branch will be dealt with before any in the other. But if there is a sizeable interval between successive balls in each branch, the door may be swung back and forth so as to deal with each branch alternately. A slow presentation of stimuli may equally allow them to be dealt with in the actual order of arrival rather than channel by channel (7, 8). It should be noted that the time allotted to "swings of the door," as measured by these experiments, agrees with that determined by the experiments on prolonged performance mentioned previously (6).

(d) Again, suppose a lengthy series of balls is inserted through one branch, the door being held open for them by the handle. If a couple of balls are meanwhile delivered to the other branch and wait there till the stem is clear for them to pass, there is naturally some risk that this extra pair of balls will exceed the critical time. The risk will be greater if they are inserted with the earlier balls of the long series than if inserted with the later balls. This also is true of immediate memory for spoken digits arriving at the two ears (9). A point of interest concerns the effect of "irrelevant" balls which arrive on the second branch during a long series on the first. If all the balls in the second branch are irrelevant, the door may be kept closed against them and they will eventually exceed the critical time and be removed (15). But if some relevant balls are also on the second branch, the door will have to be opened and the mixture of relevant and irrelevant balls passed down the stem. The former can be recirculated in the usual immediate memory fashion, and the latter removed by the lower door. The presence of irrelevant balls either before or after relevant ones will therefore produce a greater risk of jamming or of the critical time being exceeded. This also is true of immediate memory (9).

(e) Now consider cases in which the order of the balls emerging at the bottom of the stem is different from that of their insertion. The extreme case is that of "backward memory span." With our simple model, the easiest situation to consider is that in which the first half of a series of balls is to emerge after the second half. This can be arranged on the first circulation by passing the first balls down the stem and back up the return tube to the branch opposite to that in which they were inserted. The later balls are returned to the same branch which they entered, and on the second circulation are the first to be admitted to the stem. After they have passed the flap the other balls can follow them. This will mean that the first balls have stayed in the branch longer than will normally be necessary for simple recirculation, and so are more likely to exceed the critical time. Therefore such a rearrangement of order will reduce the "memory span." Furthermore it will alter the order of difficulty, since the first members of the series will suffer more than the later ones. These effects have been shown by Kay and Poulton (26), and are supported by Brown (13). The reduced memory span for a rearranged list is of course familiar from intelligence testing, but the change in the serial-order effect is a more important deduction.

(f) Serial-order effects will clearly depend very considerably on the rates of presentation of stimuli and of required response, as compared to the rate of recirculation. They will also depend on the way in which stored information is used to operate the two handles—that is, on the strategy of the subject. They will certainly not be completely determined by the primitive learning process, as is supposed by some existing theories. In fact it has been shown by Kay and Poulton (26) that the serial-order effect is altered by the subject's absence

of knowledge about the order in which recall will be required, even though the actual order remains the same as in a control experiment. (Knowledge of the amount to be recalled will also affect the efficiency of recall as well as the order, as has been shown by Brown [13]. This is because the uninformed subject must recirculate material which is not in fact to be recalled later; retention is an active process.)

Some general points about serial-order effect may be made, however. First, if the rate of response is below that of recirculation, but not so slow as to allow complete recirculation between responses, the first items in response will have stayed in store less long than the later items. So, as in (b) above, the earlier part of a list recalled in the order of presentation will be better recalled than the later part. Second, if the time taken to respond is eliminated—for instance, by requiring recall of only one item in the series—then the time taken to present the stimuli will be the chief factor influencing serial-order effects. The last stimuli to arrive will then be those which have been stored for the shortest time, and will be best recalled. This is the result of Gibson and Raffel (23).

Third, a special case of some importance arises when fresh material is being presented while earlier material is being recalled. Considering the model, suppose a short series of balls are inserted and recirculated to the branch opposite that by which they arrived. If they are now to be passed down the stem during the arrival of fresh balls on the original branch, the earliest of the second series of balls will suffer the longest delay. Consequently the end of the first series and the beginning of the second series will be the points of greatest difficulty. If the two series are considered as one long one, the familiar U-shaped serial-order effect will appear.



This has been demonstrated by Poulton (32). In conventional learning experiments subjects are not instructed to rehearse (recirculate) the earlier items during presentation of the later ones. But it seems plausible that, during the presentation of a long series at a medium rate, the presence of the earlier balls in the second branch should normally encourage a tendency to recirculate them while the later balls are still arriving. This would be comparable to the higher priority of a previously quiet channel, which was mentioned in the section on perception. It would result in the U-shaped curve of difficulty being the usual one for serially presented material; but the curve will be subject to great modification by instructions. Furthermore a very slow rate of presentation, resulting from allowing recirculation during gaps between presentations, will minimize the effect. This is the case (24).

#### LIMITATIONS OF THE MODEL

Certain properties of the model are likely to be misleading. Of these the most important has been stated several times above, but is worth repeating. The balls represent information, not stimulation. The reader must not contemplate the Y tube and decide that two stimuli cannot be dealt with simultaneously. They can if they convey sufficiently little information. Clearly, many reflexes are compatible with one another, and it is likely that simple "voluntary" reactions are equally capable of being carried on simultaneously. It is only with unpracticed reactions involving a choice between several alternatives that we find an interference between two stimuli; but this is a very normal case outside the laboratory. Theories such as that put forward by Welford (38) for the "refractory period" undoubtedly need the

qualification that highly probable stimulus sequences may not show these effects (22). But this does not disprove their general value.

A related point is more serious. The length of the immediate memory span is roughly constant whatever the size of the ensemble from which the items are chosen (31): one cannot remember enough binary digits to make the information in immediate memory equal to that stored in memorizing ordinary decimal digits (30). Yet the model makes the limit on memory span dependent on the time taken to pass the items through a limited capacity system, which will in turn depend on the information per item. Perhaps this difficulty may be resolved by suggesting that each possible item represents an extra branch on the stem, and an incoming ball is always recirculated to its appropriate branch. The time taken to withdraw all the balls again would then depend on the time taken by the flap to operate, which is not dependent on the information per item. But this is clearly leading us into complications; for the present we may merely note that the point is an important one about immediate memory but not impossible to handle with a model of this type. A minor caution which should be added is that the model is deterministic while all the experimental results quoted are statistical.

Finally, we have put forward this model as one which may be described by an exactly worded theory which applied also to man. The present paper is directed largely at those who find such a theory unintelligible in its original form, but it should be borne in mind that the theory under test is the abstract one given in the first section. Otherwise the error of identifying the model with the organism may be made, if only to discredit the theory by its obvious absurdity. The formulation given ear-



lier is not complete, but it indicates the way in which the model and the man may be described by the same abstract theory. There should therefore be no excuse for treating the Y tube as anything more than an expository device and a mnemonic for recalling the results of numerous experiments. The writer's freedom from such an error is demonstrated by the fact that he has never built his model in any physical sense.

### SUMMARY

A mechanical model is described, to act as an easy introduction to a formal theory of attention and immediate memory in information theory terms. A number of deductions from the theory which agree with experimental results on human beings are given as descriptions of the behavior of the model.

### REFERENCES

1. BARRICK, H. P., NOBLE, M., & FITTS, P. M. Extra task performance as a measure of learning a primary task. *J. exp. Psychol.*, 1954, 48, 298-302.
2. BERLYNE, D. E. Stimulus intensity and attention in relation to learning theory. *Quart. J. exp. Psychol.*, 1950, 2, 71-75.
3. BILLS, A. G. Blocking: a new principle in mental fatigue. *Amer. J. Psychol.*, 1931, 43, 230-245.
4. BROADBENT, D. E. Speaking and listening simultaneously. *J. exp. Psychol.*, 1952, 43, 267-273.
5. BROADBENT, D. E. Listening to one of two synchronous messages. *J. exp. Psychol.*, 1952, 44, 51-55.
6. BROADBENT, D. E. Noise, paced performance and vigilance tasks. *Brit. J. Psychol.*, 1953, 44, 295-303.
7. BROADBENT, D. E. The role of auditory localization in attention and memory span. *J. exp. Psychol.*, 1954, 47, 191-196.
8. BROADBENT, D. E. Successive responses to simultaneous stimuli. *Quart. J. exp. Psychol.*, 1956, 8, 145-152.
9. BROADBENT, D. E. Immediate memory and simultaneous stimuli. *Quart. J. exp. Psychol.*, 1957, 9, 1-11.
10. BROADBENT, D. E. Effects on behaviour from noises of high and low pitch. *Brit. Med. Res. Council, Appl. Psychol. Unit Rep. No. 222*, 1954.
11. BROADBENT, D. E. The concept of capacity and the theory of behaviour. In *Proc. 3rd London Symposium on Information Theory*. London: Butterworths, 1956.
12. BROADBENT, D. E. Listening between and during practised auditory distractions. *Brit. J. Psychol.*, 1956, 47, 51-60.
13. BROWN, J. The nature of set-to-learn and of intra-material interference in immediate memory. *Quart. J. exp. Psychol.*, 1954, 6, 141-148.
14. BROWN, J. Immediate memory. Unpublished doctoral thesis, Univer. of Cambridge, 1955.
15. CHERRY, E. C. Some experiments on the recognition of speech, with one and with two ears. *J. Acoust. Soc. Amer.*, 1953, 25, 975-979.
16. CONRAD, R. Speed and load stress in a sensori-motor skill. *Brit. J. indust. Med.*, 1951, 8, 1-7.
17. CONRAD, R. Missed signals in a sensori-motor skill. *J. exp. Psychol.*, 1954, 48, 1-9.
18. CONRAD, R. Adaptation to time in a sensori-motor skill. *J. exp. Psychol.*, 1955, 49, 115-121.
19. DEESE, J. Some problems in the theory of vigilance. *Psychol. Rev.*, 1955, 62, 359-368.
20. DEUTSCH, J. A new type of behaviour theory. *Brit. J. Psychol.*, 1953, 44, 304-317.
21. DEUTSCH, J. A machine with insight. *Quart. J. exp. Psychol.*, 1954, 6, 6-11.
22. ELITHORN, A., & LAWRENCE, C. Central inhibition—some refractory observations. *Quart. J. exp. Psychol.*, 1955, 7, 116-127.
23. GIBSON, J. J., & RAFFEL, G. A technique for investigating retroactive and other inhibitory effects in immediate memory. *J. gen. Psychol.*, 1936, 15, 107-116.
24. HOVLAND, C. I. Experimental studies in rote learning theory. II. Reminiscence with varying speed of syllable presentation. *J. exp. Psychol.*, 1938, 22, 338-353.
25. HYMAN, R. Stimulus information as a determinant of reaction time. *J. exp. Psychol.*, 1953, 45, 188-196.
26. KAY, H., & POULTON, E. C. Anticipation in memorizing. *Brit. J. Psychol.*, 1951, 42, 34-41.

27. MACKWORTH, N. H. Researches on the measurement of human performance. *M.R.C. Spec. Rep. No. 268*. Her Majesty's Stationery Office, 1950.
28. MACKWORTH, J. F., & MACKWORTH, N. H. The overlapping of signals for decisions. *Amer. J. Psychol.*, 1956, 69, 26-47.
29. MACKAY, D. M. In symposium on cybernetics. *Advanc. Sci.*, 1954, 40, 402-406.
30. MILLER, G. A. The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychol. Rev.*, 1956, 63, 81-97.
31. POLLACK, I. Assimilation of sequentially encoded information. *Amer. J. Psychol.*, 1953, 66, 421-435.
32. POULTON, E. C. Memorization during recall. *Brit. J. Psychol.*, 1953, 44, 173-176.
33. POULTON, E. C. Two-channel listening. *J. exp. Psychol.*, 1953, 46, 91-96.
34. POULTON, E. C. Listening to overlapping calls. *J. exp. Psychol.*, 1956, 52, 334-339.
35. VINCE, M. A. Rapid response sequences and the psychological refractory period. *Brit. J. Psychol.*, 1949, 40, 23-40.
36. SPIETH, W., CURTIS, J. F., & WEBSTER, J. C. Responding to one of two simultaneous messages. *J. Acoust. Soc. Amer.*, 1954, 26, 391-396.
37. WEBSTER, J. C., & THOMPSON, P. O. Some audio considerations in air control towers. *J. Audio. Eng. Soc.*, 1953, 1, 171-175.
38. WELFORD, A. T. The psychological refractory period and the timing of high-speed performance—a review and a theory. *Brit. J. Psychol.*, 1952, 43, 2-19.

(Received November 30, 1956)

# RIVISTA DI PSICOLOGIA

*Publication of the Italian Society of Psychology*

*Director:* PROFESSOR CESARE L. MUSATTI

*Co-directors:* PROFESSORS ALERTO MARZI, GAETANO KANIZSA, MARIO PONZO

Vol. 50, No. 4: "Jubilee" number in celebration of the fiftieth anniversary of the establishment of the Rivista.

---

## TABLE OF CONTENTS

- BARTLETT. Training Older People: An Experimental Approach to a Social Problem  
BURT. British Psychology During the Past Fifty Years  
DREVER. A Note on the Classification of Stimuli  
EYSENCK. The Questionnaire Measurement of Neurotism and Extraversion  
von FIEANDT. A Contribution Toward the Problems of Age and Personality  
GEMELLI. La percezione visiva del movimento  
GOTTSCHALDT. Handlung und Ausdruck in der Psychologie der Persönlichkeit  
de GROOT. Über das Denken des Schachspielers  
MEILI. Geistige Faktoren in den ersten Stadien der Charakterentwicklung  
METZGER. Ueber Durchsichtigkeits-Erscheinungen  
NUTTIN. Experimental and Clinical Method in Psychology  
PIAGET. Centration et décentration perceptives et représentatives  
PIÉRON. La temperature en Psychophysiologie  
PONZO. Il fattore della prospettiva in un gruppo di illusioni nelle valutazioni di collettività numeriche  
RAUSCH. Zur Theorie der sogenannten Vertikaltäuschung  
ROHRACHER. Zur Theorie der Aufmerksamkeit  
TERSTENJAK. The Experience of Time and the Question of Subjective Boredom  
WALLON. Subjectivisme et objectivité en psychologie  
WERNER WAPNER. Sensory Tonic Field Theory of Perception: Basic Concepts and Experiments  
WELLEK. Die Entwicklung der Grundannahmen der Psychologie und die Überwindung des Phänomenalismus und Psychologismus  
ZAZZO. Le bestiaire des enfants

This "jubilee" number, which will be sent gratis on 1956 subscriptions, will also be sold as a separate issue. Correspondence and subscriptions should be addressed to:

Amministrazione Rivista di Psicologia

Via degli Alfani 37/39

Firenze, Italia



# THE PSYCHOLOGICAL REVIEW

## DISCRIMINATION OF CUES IN MAZES: A RESOLUTION OF THE "PLACE-VS.-RESPONSE" QUESTION<sup>1</sup>

FRANK RESTLE<sup>2</sup>

*Center for Advanced Study in the Behavioral Sciences*

Whether rats in mazes learn turning responses or places is a question which has often been subjected to experimental test. Data from different experiments conflict, and attempts at a definitive answer seem only to add to the confusion. This paper will defend the thesis that the place-vs.-response question is wrongly formulated, and that the data which fail to decide between place and response learning give an unequivocal answer to a question properly stated.

The place-vs.-response question has been approached largely through the use of the T maze, rotated in its visual surround. A typical arrangement showing runs reinforced in "place" and "response" learning is shown in Fig. 1, along with the fixed-maze problem in which both place and response may be learned. In both place learning and response learning, the maze is rotated on alternate trials at random. The place learner is always to go to the same place in the room, responding consistently to extra-maze cues but making different turns on different trials. The response learner makes the same turn on all trials, going to different places. One supposed test of whether place or response learn-

ing is more dominant is to compare rates of learning on these two problems. Another test is to train animals with the maze in a fixed position (place + response learning), and then rotate the maze for a test trial. The animal can now either make the same turn he has learned or go to the same place he has been going to, but not both. In this *direct opposition* experiment, the relative number of animals taking each choice is a test of the relative dominance of place and response.

In 1946, Tolman, Ritchie, and Kalish (23) proposed that in such a situation place learning is more natural and primitive for the rat than response learning, and place will dominate response in all tests. Their experiment supported the hypothesis, but some later repetitions have found response dominating place, or have found no difference.

An answer to these apparent contradictions is here sought by assuming that maze running depends on a multiplicity of cues, and that the rat learns differential responses to relevant cues in a maze just as he would in a discrimination box. A theory based on learning in the discrimination box will be applied to the maze situation. Since the discrimination theory to be used (15) is very close to earlier theories of the effect of sensory input on learning in mazes, theories which antedate the place-vs.-re-

<sup>1</sup> Dr. Richard L. Solomon suggested and drafted part of this paper (see footnote 3). His guidance and help are gratefully acknowledged.

<sup>2</sup> Now at Michigan State University.

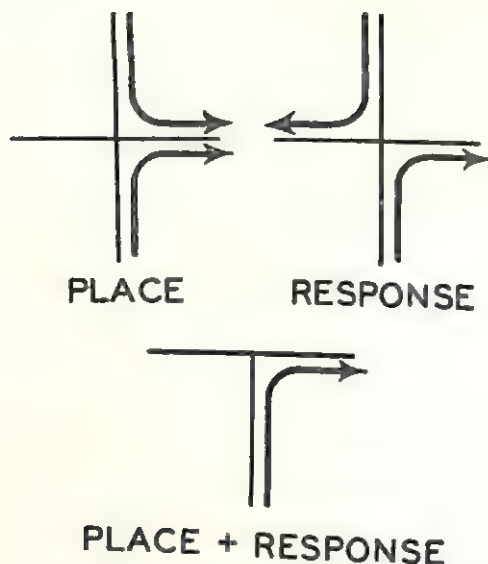


FIG. 1. Runs reinforced in place, response, and place + response (fixed maze) learning. Only one T maze is present on a given trial.

sponse controversy, it is useful to place the recent studies within their historical context and analyze all the data at once.

#### MULTIPLE CUE THEORY

Hunter (9, 10) and Honzik (8), among others, have proposed that learning and performance in mazes depends upon many cues, in all available modalities and from all sources. Depending on what stimuli are available to the rat, maze learning may depend on visual, auditory, olfactory, tactual, and kinesthetic stimuli. Visual, auditory, and olfactory stimuli may arise from within the maze or outside. The cues used by rats may be not only the specific physical stimuli but also patterns or arrangements of stimuli, so that changes which preserve the arrangement may not disrupt performance.

When a portion of the cues are removed by surgery or screening, or made irrelevant by interchanging maze units, rotating the maze, etc., learning of the maze is retarded, but mazes are often perfectly solved in the presence of known irrelevant cues.

The earlier formulation of multiple cue theory is somewhat refined by considering more recent theory derived from discrimination learning (15). This newer theory distinguishes between relevant cues, which bear a regular relationship to the correct path, and irrelevant cues which are not consistently related to the correct response.

A cue with a constant or predictable relationship to the true path will in the course of learning become "conditioned" to the correct response. All such cues will eventually be learned, and all will play a part in performance. If a certain cue is irrelevant, bearing a changing and unpredictable relationship to the correct response, it cannot be the basis of learning and is only a distractor. Such cues will become "adapted" during learning, and eventually will play no part in performance.

The rate of learning depends directly on the proportion of relevant, usable cues in the total set available.

When learning has been based on a variety of relevant cues and then some of these cues are scrambled and made irrelevant, the amount of disturbance reflects the relative importance of the newly changed cues. Though irrelevant cues will eventually be adapted, making important cues irrelevant reduces the proportion of relevant cues and thus retards learning. Disruption due to scrambling of learned relevant cues is attributed to the fact that such cues are not adapted, having previously been relevant. Recovery from the disturbance results from progressive adaptation of the scrambled cues.

If the learning program involves constantly introducing new cues which have not been present before, these new cues will always disturb performance, since they cannot have been attached to the correct response and they cannot have been adapted.

In an ordinary multiple-unit maze, the untrained rat responds erratically early in training, producing different kinesthetic stimuli on different trials. As a result, such stimuli are not consistently related to the true path, and alone they cannot be the basis of rapid learning. As performance based on other cues improves, the animal makes more nearly the same movement each trial, producing for himself a regular pattern of kinesthetic stimuli which can serve as partial basis for the learned habit. Highly skilled performance thus depends to some degree on kinesthesia, though such cues do not alone mediate learning to a significant degree.

The above application of the theory of discrimination learning to mazes does not differ in important respects from the assumptions and conclusions of Hunter and Honzik. The main refinement is in attributing the rate of learning to the *proportion* of relevant cues. The earlier writers do not make just this statement, though their statements may be interpreted to this effect.

No attempt will be made in this paper to apply the specific quantitative formulation of discrimination learning to multiple-unit maze learning. The effects peculiar to the *serial* character of maze learning are not reflected in discrimination theory. When a single-unit T maze is used, as in the recent place-response experiments, discrimination theory should apply exactly and quantitative predictions should be correct.

The multiple cue theory can be compared with the results of studies using multiple-unit mazes, with the understanding that specific quantitative values (error scores, trials to learn, etc.) depend both on the discrimination of cues and on serial patterns. An understanding of these earlier studies of maze cues is essential in gaining a clear insight into the place-vs.-response controversy.

### THE HISTORICAL CONTEXT<sup>3</sup>

Studies of the sensory control of maze behavior of rats have been reviewed in detail by Munn (14), so only a broad summary will be given here, along with discussions of the main points.

Early studies of the role of sensory processes in maze behavior were motivated by the early conclusion by Small (1901) and by Carr and Watson (1908) that in complex mazes rats form kinesthetically controlled habits which become "automatic." This hypothesis is that kinesthetic cues, arising from one response in the maze, serve as the main cues to the next response in the sequence. Each response is associated with the previous one, and other sensory input becomes unimportant in maintaining performance. This conclusion was based on the erroneous assumption that if one sort of cue is important, others must be unimportant. The technique was essentially one of eliminating a single sense modality at a time, in each case observing that performance remains essentially intact. The only modality not disturbed was kinesthesia, which was assumed to be "the" crucial one. Hunter (11) has shown the flaw in this line of argument.

The "kinesthetic" hypothesis inspired a series of experiments designed to prove or disprove it. These experiments uncovered a wealth of information about the cues which in fact control behavior in various mazes. Fundamentally, the data were analyzed to determine the relative effectiveness of (a) kinesthetic stimuli, (b) intra-maze stimuli, such as visual, olfactory, and tactual stimuli from the maze itself, and (c) extra-maze stimuli, such as visual or auditory

<sup>3</sup> The bearing of these studies on the place-vs.-response question was pointed out to the author by Dr. R. L. Solomon. This section and the beginning of the next section are based in part on a personal communication from him.



stimuli from the room containing the maze. Intra- and extra-maze stimuli were broken down into components due to different sense modalities, again with the intention of evaluating each.

The methods used in assessing the role of sensory events in maze behavior were as follows: (a) surgical interference with receptor organs and neural pathways; (b) elimination of stimuli from the intra-maze or extra-maze environment; (c) introduction of distinctive stimuli to the environment; and (d) the controlled rearrangement of intra-maze and extra-maze stimuli.

In general, the results of these experiments were as follows:

(1) Simple alternation mazes can be learned by the rat on the basis of kinesthetic cues alone, but more complex mazes cannot be learned in any reasonable number of trials without the aid of visual, olfactory, or auditory cues. Removal of kinesthetic cues by surgical means does not greatly affect maze performance if intra- or extra-maze cues are left intact. However, once a maze performance is perfected on the basis of intra- or extra-maze cues, removal of such cues does not destroy the performance completely. As one would expect from the multiple cue theory, kinesthetic cues are relevant and conditioned to the correct response only when a regular relationship between the last response and the next correct response is established. Such a relationship exists in a simple alternation maze early enough for learning to take place. In more complex mazes, usable kinesthetic cues exist only after the maze is learned on some other basis. Since during learning kinesthetic cues are mostly irrelevant anyway, their removal does not retard learning.

(2) Mazes can ordinarily be learned on the basis of intra-maze cues alone, even if extra-maze cues are made irrelevant by rotating the maze in the room. Such rotation retards learning, however,

especially if the maze is elevated and the room contains conspicuous visual cues. If the maze is kept in place in the room and its spatial arrangement is preserved while units of the maze are interchanged to make intra-maze stimuli irrelevant, rats can still solve the problem. Again, interchanging of stimuli retards learning. In many mazes, it should be noted, all blinds have a common visual appearance, being shorter than true alleys and ending within sight of the rat at the choice point. Interchanging units does not make such stimuli irrelevant. But even if the spatial pattern of the maze is scrambled, if the goal is in a fixed position relative to the extra-maze environment the rat can learn to run to the goal box without following alleys which lead away from the food. Thus, intra-maze and extra-maze cues are each separately capable of sustaining learning and performance in the maze. With one type scrambled, perfect performance can often be attained based on the other type, indicating that irrelevant cues are eventually disregarded. Retardation of learning due to removal or scrambling of important cues is consistently observed.

(3) In elevated mazes, where most of the visual field arises from outside the maze, extra-maze cues are usually more important than intra-maze cues. If the maze is enclosed in a homogeneous room, however, extra-maze cues are relegated to a minor role. Intra-maze cues are generally more important in alley or tunnel mazes, where the rat has at best an obstructed view of the outside. In a unidirectional maze, extra-maze cues are more important than they are in a maze which requires the rat to run in many different directions. The results support the idea that either type of cue may be the more important, depending on the relative amount of relevant stimulation stemming from each source.

(4) In general, two ways of assessing

the importance of a certain cue give comparable estimates. One method is to scramble the cue during learning, assessing its importance by the relative retardation in the rate of learning. The other method is to have the rat learn the maze with the cue relevant, then scramble it. The amount of disruption of the perfected habit indicates the importance of the cue. To an approximation, these two methods rank various classes of cues the same way. Removing a type of cue by depriving the animal of necessary receptors, as by blinding or deafening the animal or rendering it anosmic by surgical interference, is not the same thing as scrambling cues experimentally. Blinding, for example, removes both relevant and irrelevant visual cues, whereas rotating the maze makes otherwise relevant extra-maze cues irrelevant, and does not affect intra-maze visual cues. Thus, exact comparisons between the effects of scrambling and the effects of surgery cannot usually be made.

These findings suggest that kinesthetic cues are unimportant in the learning of complex mazes, but that both intra- and extra-maze cues are important. The relative importance of various types of cues depends on the maze and its surround, for rats seem to use various cues proportionally as those cues are available and relevant. Rats seem capable of overcoming the distraction of irrelevant cues.

The results are consistent with the conception that maze learning and performance depend on multiple cues, and that such cues are discriminated and responses learned to them in accord with the theory of discrimination learning.

#### "PLACE" AND "RESPONSE" IN SINGLE-UNIT T MAZES

In the light of the earlier extensive studies of the sensory basis of maze run-

ning, studies of place and response in T mazes appear as comparisons of extra-maze and kinesthetic cues. The Tolman-Ritchie-Kalish hypothesis that rats learn "places" rather than "responses" means, in the earlier terminology, that extra-maze cues are more important than kinesthetic cues. Since in earlier studies kinesthetic cues were found barely sufficient to permit any learning at all, whereas extra-maze visual cues bulk large in importance when they are available, it would appear that the dominance of "place" learning was indubitable, and the Tolman-Ritchie-Kalish experiment redundant.

It should be noted that Tolman's "place" formulation is inferior to earlier formulations. Tolman does not specify what cues are thought to make up "place" indications, and he does not identify "response" learning as learning based solely on kinesthetic cues. Though intra-maze cues do not seem to be "response" cues in Tolman's sense, it is not clear whether they are "place" cues or not.

Though the hypothesis that place learning will dominate response learning seems to follow in a general way from Tolman's concept of cognitive maps, his position was not really controversial because stimulus-response theorists did not believe response learning would be easier than place learning. The early Carr-Watson hypothesis, that maze habits are mainly controlled by kinesthesia, was by 1946 buried under conflicting evidence, and S-R theories stated that the stimulus components of the maze habits were quite likely to be extra-maze visual cues, especially if the maze is elevated and in a room full of such cues.

From these considerations, one should expect that the dominance of place learning is a foregone conclusion. The existence of a "controversy" is itself a



surprise. In order to show the nature of this controversy it is necessary first to consider how place-response experiments have been conducted and what results have been obtained. Since this recent literature has not received a definitive review, a relatively thorough analysis is presented here.

### *Tests of Place-vs.-Response Dominance*

We first consider whether place learning in fact dominates response learning in the single-unit maze. Ten studies have compared the rates of place and response learning in single-unit T mazes or slight modifications thereof. Of these, seven found place learning faster (1, 3, 5, 22, 23, 24, 25), and two found response learning faster (7, 20), while one found that either could be faster depending on the intertrial interval (21). In direct opposition tests (rotation of the maze after the habit is learned on a fixed maze), one study showed that either place or response could be superior depending on the shape of the approach stem (18), another showed that either could be dominant depending on differential cues (27), and two others found response tendencies overriding place tendencies (2, 12). A modification of the opposition test consists of setting up conditions in which the animal alternates. Several studies have asked whether an animal alternates response or stimulus (place) characteristics (4, 6, 13, 26), and the answer has always been that place alternation is stronger than response alternation. These studies used enclosed alley mazes, "place" cues being especially distinctive intra-maze cues.

In summary, if we merely count titles, the impression is received that place tendencies are usually learned faster than responses, and that they are sometimes stronger in opposition tests, especially those depending on alternation.

One certainly cannot draw the conclusion that place learning is always dominant, though, for response tendencies dominate in at least some conditions of seven different studies.

### *Conditions Associated With Place and Response Dominance*

Since neither place nor response is uniformly dominant, we may consider the experimental conditions which make one or the other stronger.

Several writers (1, 7, 21, 25) have suggested that the relative dominance of place depends on the amount of differential visual stimulation. This is, of course, the position taken by earlier writers on general maze learning, though place-vs.-response studies are not usually analyzed in terms of this variable. Counting heads in the experimental reports, we see that the use of a homogeneous visual surround (a dome or enclosure, usually made of muslin, which prevents the animal from discerning any uncontrolled stimuli from the room about him) greatly predisposes rats to learn responses instead of places. All four experiments using such domes report response dominance (2, 7, 20, 21), though one (21) showed that response dominance could be neutralized by massing trials. The only other cases of response dominance reported (12, 18) were in plain rooms under low illumination, which may be thought of as approximations to dome-type enclosures. All other studies showing place dominance were done in open rooms (1, 3, 5, 22, 23, 24, 25) or in alley mazes with strong differential cues (6, 13, 26). One study showed that when the two places contrast sharply in illumination, place dominates, whereas when the illumination is more nearly equal, response tendencies dominate.

In summary, place tendencies dominate when visual stimuli at the two ends



of the maze are very *unlike*, and response tendencies dominate when such stimuli are relatively *alike*. In every case the domination seems to be quantitative—both place and response tendencies exist, but one is stronger than the other depending on stimulus conditions.

By inspecting the experimental reports we can get some idea of what constitutes a strong place cue for the rat. The most dramatic place dominance was found in two studies (17, 23) in which rat cages were located to one side, nearer one goal than the other. In one of these studies (17) it was found that the rats would not give up responding to one place in the room, despite a number of controls, until the cages were moved, following which almost all the rats reversed. Stimuli such as lights had only a slight effect on performance, compared with the rat cages.

When rat cages are placed directly behind the starting point and are thus not available as place cues, or when the room is devoid of cages, such visual stimuli as windows (in daylight studies) give rise to strong place preferences (1, 3), as do well-lighted rooms with many small objects in them (5, 25). Somewhat less striking but still consistent place dominance was shown in two studies in which one wall of a plain room (the wall behind one goal), was moderately well illuminated, the other wall not being illuminated at all (22, 24). The power of room cues can be estimated by noting the quick learning of a successive discrimination between rooms (19). Attempts to give differential cues inside domes by illuminating a 10-inch disc behind one goal with a 7.5-watt lamp (2, 12) or using 7.5-watt lamps themselves as cues, with (7) or without (20) overhead illumination, did not lead to place dominance. Symmetrical overhead illumination in an empty room produced response dominance with spaced practice

(21). A small lamp on the floor below one goal, pointed to throw long shadows on the floor, gave slightly more place than response tendency (18). Ratio of illuminations rather than difference seems to be a crucial variable (27).

In summary, rat cages are very strong place stimuli, windows or objects in a well illuminated room rank next in power, and fairly strong differential lighting of walls of a plain room is also effective. Discs lit by 7.5-watt lamps, or such lamps themselves, are not usually enough to make place dominant over response in an otherwise homogeneous setting.

Some other variables are associated to some degree with place and response dominance. For example, all experiments using pigmented rats, (4, 22, 23, 24) show place dominance, except for one (12) which showed cases of both place and slight response dominance under low illumination. But white rats also show place dominance under good illumination (1, 3, 5, 25), so we may reason that it just happened that experimenters who intended to use strong visual cues also sometimes used pigmented rats to take advantage of their superior vision. The pigmentation of the rat cannot be shown to have any other effect on place and response dominance.

The noncorrection procedure (picking up the rat after an error) seems to lead to better response learning than the correction procedure (allowing the rat to find the food by retracing after an error). Using the place-response studies, the effects of correction cannot be isolated from the visual cue effects mentioned above.

One study (21) indicates that response learning is slowed by massing trials. Little more about massing effects in these studies is known, but most

of the studies employ at least moderate trial spacing.

There seems to be no interesting relation between place-vs.-response dominance and either the ages of rats used or the size of the maze employed. The slight tendency for smaller mazes to go with response dominance is more than explained by the tendency of experimenters using domes to use slightly smaller mazes for convenience.

### *Evaluation of the Place-vs.-Response Controversy*

The details of the single-unit T maze experiments quite clearly indicate that there is nothing in the nature of a rat which makes it a "place" or "response" learner. The main factor determining the outcome of place-vs.-response experiments is the amount of extra-maze visual stimulation which differentiates the region around and behind one goal from the region around and behind the other. Such visual cues are relevant in place learning and irrelevant in response learning.

It seems reasonable to conclude that the place-vs.-response controversy, which seemed ill formulated when compared above with earlier maze studies, gives a distorted and confusing interpretation of the experiments designed to settle it.

It should be remarked that in single-unit mazes, kinesthetic cues appear sufficient to support quite rapid learning. This conflicts with Honzik's conclusion that "when all other avenues of stimulation are destroyed, kinaesthesia is helpless" (8, p. 56). Honzik's statement, if taken as applying generally to all mazes including single-unit ones, would be incorrect. The importance of a sense modality depends on the richness and relevance of stimuli in that modality which exist in the maze situation. In single-unit mazes, relevant kinesthetic stimuli apparently abound. In Honzik's 14-

unit maze, animals do not make regular enough runs to give themselves a constant set of kinesthetic stimuli on which to build accurate performance. One may also consider Hunter's point that in complex mazes, with left and right turns required in irregular order, simple kinesthetic cues are irrelevant. Attempts to state the relative importance of sense modalities or of intra- and extra-maze cues in general, for all mazes, are akin to the "place-vs.-response" hypothesis in that they fail to incorporate the most important variable, the stimulus situation presented to the animal. Such attempts are, accordingly, doomed to failure.

### PREDICTION OF T-MAZE DATA BY THE QUANTITATIVE THEORY OF DISCRIMINATION LEARNING

The single-unit T maze, being devoid of the serial characteristics of more complex mazes, may be thought of as a kind of discrimination-learning apparatus. When the T maze is rotated on random trials as in place-response experiments, extra-maze (place) and kinesthetic (response) cues are uncorrelated.<sup>4</sup> In place learning (see Fig. 1) place cues are relevant, and all others including response cues are irrelevant. In response learning, response cues are relevant and all others including place cues are irrelevant. In place + response learning, with a fixed maze, both place and response cues are relevant.

We may entertain the hypothesis of cue-additivity: that the set of cues rele-

<sup>4</sup> If the whole maze is rotated as a unit, intra-maze cues are relevant in response learning and are confounded with kinesthetic cues. If the starting stem is moved but the cross-arm is left in place, intra-maze cues are relevant in place learning, and are confounded with extra-maze cues. We shall assume that intra-maze cues combine additively with the cues they are correlated with, and shall not specify what happens to them.

vant in place + response learning is simply the sum of the place and response cues. This hypothesis can be tested quantitatively, using a theory of discrimination learning (15). In this theory, the rate of learning is set equal to the proportion of relevant cues, and the single number is called  $\theta$ . The learning curve (or statistics such as total errors to mastery, total errors in  $n$  trials, trials-to-criterion, etc.) may be used to estimate  $\theta$ : and, conversely, given  $\theta$  the learning curve or any of its statistics can be computed in advance of experimentation. The equations involved and the methods of computation will not be repeated here.

Galanter and Shaw (5) used three groups of rats trained in the same apparatus and surround, under conditions of place learning, response learning, and fixed-maze or place + response learning, respectively. Using median trials-to-criterion<sup>5</sup> as an index of typical performance, the proportion of relevant cues in each problem was computed, using the equations of discrimination-learning theory. Theoretically, the proportion of relevant cues in the place + response problem should be the sum of the proportions in the other problems:  $\theta_{P+R} = \theta_P + \theta_R$ . The results of the computations to check this hypothesis are shown in Table 1. Place + response performance is predicted using only data from the place and response groups. Inspection of Table 1 indicates that the prediction is relatively accurate, although, since subjects in the place + response

TABLE 1

MEDIAN TRIALS TO 10-OUT-OF-10 CRITERION  
AND PROPORTION OF RELEVANT CUES ( $\theta$ )  
IN THE GALANTER-SHAW EXPERIMENT

Relevant Cues	Observed		Predicted	
	Trials	$\theta$	Trials	$\theta$
Place	5.0	.34	—	—
Response	33.5	.10	—	—
Place + Response	2.0	.58	3.1	.44

group ranged in trials-to-criterion from 0 to 3, the prediction is at the edge of the obtained distribution of scores.<sup>6</sup>

Scharlock (20) used a maze in a dome with either one light behind one goal or lights behind both. If there was one light it was the only source of differential place cues, so if there were two lights there were no place cues. Scharlock ran place, response, and place + response groups with one light. He also had a control group which made no progress on place learning with both lights—an expected result confirming that with a light behind each goal there were no place cues. One other group, which we may call “response-minus-place,” learned a response with both lights on, and thus with no place cues available. Here again we can predict that the place + response group will yield a learning rate,  $\theta_{P+R}$ , which is the sum of the rates of the place and response groups,  $\theta_P + \theta_R$ . In addition, the learning rate of the response-minus-place group should be faster than that of the response group, because of the

<sup>5</sup> Since retracing and “false-start” responses were counted as errors, animals had a higher probability of making an error than a correct response at the beginning of training. This bias in favor of errors has relatively little effect on trials-to-criterion, which is accordingly used as an index of performance. The parameter  $\theta$  is estimated from trials-to-criterion by an approximation to the maximum likelihood method.

<sup>6</sup> Dr. Galanter, in a personal communication, noted that the place + response group was run after the other parts of the experiment were completed. In Galanter's opinion, the experimenters were by this time somewhat more skillful, and the place + response group had an advantage. This might account for the discrepancy between prediction and observation.



TABLE 2

ERRORS IN 28 TRIALS AND PROPORTION OF  
RELEVANT CUES ( $\theta$ ) IN THE SCHARLOCK  
EXPERIMENT

Relevant Cues	Observed		Predicted	
	Errors	$\theta$	Errors	$\theta$
Place	5.53	.216*	—	—
Response	3.84	.296	—	—
Place learning, no place cues (2 lights)	17.50	.000	—	—
Place + Response	2.28	.445	2.00	.512
Response - Place: no place cues (2 lights)	2.84	.366	2.66	.378]

\* Because of initial biases, the probability of correct response on the first trial is estimated at about .375 for all groups. The correction is made in the fashion shown in (16).

elimination of irrelevant place cues. Since  $\theta_P$  is the proportion of differential place cues, it follows theoretically that  $\theta_{R-P} = \theta_R / (1 - \theta_P)$ . The results of computations to check these hypotheses are shown in Table 2. Inspection of Table 2 shows that the predictions are quite accurate. The discrepancies between predictions and observations are not statistically significant.

A third experiment by Blodgett, McCutchan, and Mathews (3) separates location and direction cues. The rat may approach the same location (for instance, the center of the room) from either of two directions if the maze is shifted appropriately. Location and direction are usually lumped as place cues, but in this experiment they are separated. Seven groups constitute the experiment, with one group learning each problem possible: location, direction, response, all combinations of two relevant, and the combination of all three relevant (fixed maze learning).

The data were reported in terms of "cycles," pairs of trials, which contained at least one error. From this it is not possible to make good estimates of  $\theta$ , but an effort has been made to attain fair approximations. The estimates, and the corresponding predictions made by

adding  $\theta$ -values of problems with fewer relevant cues, are shown in Table 3. The results seem quite encouraging, under the circumstances.

### SUMMARY AND CONCLUSIONS

Consideration of early studies of the sensory basis of maze learning, and review of place-vs.-response experiments, indicate that:

(1) There is nothing in the nature of a rat which makes it a "place" learner, or a "response" learner. A rat in a maze will use all relevant cues, and the importance of any class of cues depends on the amount of relevant stimulation provided as well as the sensory capacities of the animal. In place-response experiments, the importance of place cues depends on the amount of differential extra-maze stimulation.

(2) A multiple-cue theory of maze learning is successful in comprehending the major results of experiments using complex mazes, and the detailed results of place-response experiments using single-unit T mazes.

(3) Useful refinements of classical multiple-cue theory were taken from discrimination-learning theory. These are that irrelevant cues are adapted during learning, and that the rate of

TABLE 3  
PROPORTION OF RELEVANT CUES ( $\theta$ )  
ESTIMATED AND PREDICTED IN  
THE BLODGETT, MCCUTCHAN,  
AND MATHEWS EXPERIMENT

Relevant Cues	Estimated $\theta$	Predicted $\theta$
Location	.02	—
Direction	.11	—
Response	.08	—
Location + Direction	.13	.13
Location + Response	.10	.10
Direction + Response	.13	.18
Location + Direction + Response	.18	.21

learning depends on the *proportion* of relevant cues.

(4) Quantitative analysis of the results of certain place-response experiments indicates that place and response cues combine additively in the place + response (fixed maze) problem.

The writer's general conclusion is that further "definitive" studies of the place-vs.-response controversy, to prove that rats are by nature either place or response learners, would be fruitless since the issue is incorrectly drawn. However, use of the T maze to analyze the stimuli in maze learning holds promise of yielding a consistent quantitative account of how rats find their way. Such studies can build on the earlier work on more complex mazes.

## REFERENCES

1. BLODGETT, H. C., & McCUTCHAN, K. Place versus response learning in the simple T-maze. *J. exp. Psychol.*, 1947, 37, 412-422.
2. BLODGETT, H. C., & McCUTCHAN, K. Relative strength of place and response learning in the T maze. *J. comp. physiol. Psychol.*, 1948, 41, 17-24.
3. BLODGETT, H. C., McCUTCHAN, K., & MATHEWS, R. Spatial learning in the T-maze: the influence of direction, turn, and food location. *J. exp. Psychol.*, 1949, 39, 800-809.
4. GALANTER, E. H. Place and response learning: learning to alternate. *J. comp. physiol. Psychol.*, 1955, 49, 17-18.
5. GALANTER, E. H., & SHAW, W. A. "Cue" vs. "reactive inhibition" in place and response learning. *J. comp. physiol. Psychol.*, 1954, 47, 395-398.
6. GLANZER, M. The role of stimulus satiation in response alternation. *J. exp. Psychol.*, 1953, 45, 387-393.
7. HILL, C. W., & THUNE, L. E. Place and response learning in the white rat under simplified and mutually isolated conditions. *J. exp. Psychol.*, 1952, 43, 289-297.
8. HONZIK, C. H. The sensory basis of maze learning in rats. *Comp. Psychol. Monogr.*, 1936, 13, 1-113.
9. HUNTER, W. S. The sensory control of the maze habit in the white rat. *J. genet. Psychol.*, 1929, 36, 505-537.
10. HUNTER, W. S. A further consideration of the sensory control of the maze habit in the white rat. *J. genet. Psychol.*, 1930, 38, 3-19.
11. HUNTER, W. S. A consideration of Lashley's theory of the equipotentiality of cerebral action. *J. gen. Psychol.*, 1930, 3, 455-468.
12. McCUTCHAN, K., RETHLINGSHAFFER, D., & NICHOLS, J. W. The role of response and place learning under alternating hunger and thirst drives. *J. comp. physiol. Psychol.*, 1951, 44, 269-275.
13. MONTGOMERY, K. C. A test of two explanations of spontaneous alternation. *J. comp. physiol. Psychol.*, 1952, 45, 287-293.
14. MUNN, N. L. *Handbook of psychological research on the rat*. New York: Houghton Mifflin, 1950.
15. RESTLE, F. A theory of discrimination learning. *Psychol. Rev.*, 1955, 62, 11-19.
16. RESTLE, F. Theory of selective learning with probable reinforcements. *Psychol. Rev.*, 1957, 64, 182-191.
17. RITCHIE, B. F. Studies in spatial learning: III. Two paths to the same location and two paths to different locations. *J. exp. Psychol.*, 1947, 37, 25-38.
18. RITCHIE, B. F., AESCHLIMAN, B., & PEIRCE, P. Studies in spatial learning: VIII. Place performance and the acquisition of place dispositions. *J. comp. physiol. Psychol.*, 1950, 43, 73-85.
19. RITCHIE, B. F., HAY, A., & HARE, R. Studies in spatial learning: IX. A dispositional analysis of response performance. *J. comp. physiol. Psychol.*, 1951, 44, 442-449.
20. SCHARLOCK, D. P. The role of extramaze cues in place and response learning. *J. exp. Psychol.*, 1955, 50, 249-254.
21. THOMPSON, M. E., & THOMPSON, J. P. Reactive inhibition as a factor in maze learning: II. The role of reactive inhibition in studies of place learning versus response learning. *J. exp. Psychol.*, 1949, 39, 883-891.
22. TOLMAN, E. C., & GLEITMAN, H. Studies in spatial learning: VII. Place and response learning under different conditions of motivation. *J. exp. Psychol.*, 1949, 39, 653-659.

23. TOLMAN, E. C., RITCHIE, B. F., & KALISH, D. Studies in spatial learning: II. Place learning versus response learning. *J. exp. Psychol.*, 1946, 35, 221-229.
24. TOLMAN, E. C., RITCHIE, B. F., & KALISH, D. Studies in spatial learning: V. Response learning vs. place learning by the non-correction method. *J. exp. Psychol.*, 1947, 37, 285-292.
25. WADDEL, D., GANS, S., KEMPNER, P., & WILLIAMS, A. A comparison of place and response learning in very young rats. *J. comp. physiol. Psychol.*, 1955, 48, 375-377.
26. WALKER, E. L., DEMBER, W. N., EARL, R. W., & KAROLY, A. J. Choice alternation: I. Stimulus vs. place vs. response. *J. comp. physiol. Psychol.*, 1955, 48, 19-23.
27. WEBB, W. B. A study in place and response learning as a discrimination behavior. *J. comp. physiol. Psychol.*, 1951, 44, 263-268.

(Received September 21, 1956)



# LATENT EXTINCTION AND THE FRACTIONAL ANTICIPATORY RESPONSE MECHANISM<sup>1</sup>

HOWARD MOLTZ

*Brooklyn College*

## I

Most students of learning would agree that in order for an organism to acquire an adaptive behavior pattern not only must certain appropriate response be learned, but other inappropriate or non-adaptive ones must be extinguished.<sup>2</sup> It is not surprising, therefore, that the attempt to provide an explanation of the manner in which response extinction occurs is at present a strategic area of concern for current theories of learning. Of importance to such theories is a procedure, employed recently in several studies (11, 32, 36, 38), that was designed to demonstrate that the strength of an instrumental response can be weakened prior to its first nonrewarded performance. This procedure involves placing experimental animals (rats), following the termination of training, directly into a now empty but previously baited goal box for a period judged sufficient to "inform" them of its present nonrewarding character. The effect upon response strength was observed when the subsequent performance of these Ss in approaching the empty goal location was compared with that of control animals which had not received this pre-extinction experience. The pro-

cedure whereby an animal is introduced directly into an unbaited goal location after having acquired a response instrumental in securing a reward object previously contained in that location will be referred to as "latent extinction." While some investigators have questioned the effectiveness of this procedure (9, 37), others (10, 11, 23, 29, 32, 38) have demonstrated conclusively, in several different learning situations, that latent extinction does produce a decrement in the strength of the response to be eliminated. The fact that latent extinction has proven effective in this respect has been considered embarrassing to the S-R reinforcement theorist, since the implication appears to be that "... all that seems necessary for extinction is for the organism to *perceive* the absence of reinforcement" (12, p. 62). It is the purpose of the present paper to attempt to interpret the effects of latent extinction in a manner consistent with an S-R theoretical approach and, on the basis of this interpretation, to suggest hypotheses for further empirical study. It should be emphasized that the term "latent extinction" will be used throughout to denote a particular experimental procedure and *not* a process or a state of the organism.

Latent extinction has been employed primarily in two essentially different situations. The situation we will hereafter refer to as Type I was designed to measure the effect of latent extinction on the strength of a *previously established* instrumental response. The Type I situation involves training an animal to traverse a straight alley or acquire a position habit in a maze in order to ob-

<sup>1</sup> The writer wishes to acknowledge with appreciation the helpful comments of his colleagues, Professors Wayne Dennis, Elizabeth Fehrer, and David Raab.

<sup>2</sup> Extinction will be defined as the reduction in response strength that occurs following non-reinforcement. The mechanism mediating this reduction is left open, since in the opinion of the present author much more data are required before a complete and empirically adequate explanation of the extinction process can be formulated.

tain a reward object. Following training, the animal is placed directly into the goal location in the absence of the reward object for a predetermined period (latent extinction) after which "regular" response extinction is administered. The effect of latent extinction in reducing the strength of the instrumental response is indexed by the number of trials required to reach an extinction criterion. An experiment reported recently by Deese (11) serves to illustrate the Type I procedure. Deese trained rats to a position habit in a U maze with food as the reward. Following the termination of training, the experimental Ss were placed in the now empty goal location for four 1-min. periods. When these Ss were subsequently run to extinction, they yielded a significantly smaller proportion of correct responses than control Ss that had not received the latent extinction experience. Deese's results indicate that a response *can* be weakened by placing S in the goal location without the presence of the reward object and that therefore, in at least some learning situations, performance is *not* a necessary condition for extinction.

The procedure that we will hereafter refer to as Type II is concerned with the effect of latent extinction on the capacity of secondary reward stimuli to mediate the learning of a *new response* in the absence of primary reinforcement. The training employed in the Type II situation involves the use of differential reinforcement in a straight alley. By means of this technique, the cues in the baited goal box come to acquire secondary reward value while those of an unbaited goal box of a different color do not. Following the termination of training, the animal is given latent extinction which, as in the Type I situation, consists of introducing him directly into the positive goal box in the absence of the reward object. The two goal boxes

are then placed on a T maze and the number of choices to the side of the maze containing the positive goal box is taken as an index of the effect of latent extinction on the capacity of secondary reward stimuli to mediate new learning. The results of a recent experiment by Moltz and Maddi (33), to be discussed in detail below, attest to the efficacy of latent extinction in reducing the acquired reward value of previously neutral stimuli.

## II

An expectancy theory of extinction (44) encounters little difficulty in providing an interpretation of the effects of latent extinction, since it places no special requirement on the response to be eliminated. All that appears necessary to decrease an animal's readiness to respond in accordance with an established expectancy is to change the relationship between environmental objects, and to allow the animal to experience the new relationship on one or more occasions. Since the acquisition of a new cognitive pattern only requires that S, when motivated (45), experience stimulus events in spatial and temporal contiguity, response performance as such is not made a necessary component of the extinction process. Thus placing S directly into the goal location, and permitting him to observe that the reward object is no longer present, should weaken his previously acquired cognition concerning the character of the goal location. As a consequence, he will show a decreased tendency to respond in accordance with this cognition.

Hull's inhibition theory of extinction (21), on the other hand, encounters a great deal of difficulty in providing an explanation of the effects of latent extinction (especially when latent extinction is employed within the framework of a Type I situation), since the performance of the response to be elimi-



nated is considered a necessary component of the extinction process. Reactive inhibition ( $IR$ ) and conditioned inhibition ( $S^I R$ ), which are assumed to produce extinction by jointly opposing reaction potential ( $S^E R$ ), develop only in the course of response performance. This implies that either the response to be eliminated or some other response from which extinction effects can generalize must be performed before extinction can take place. Since the latent extinction procedure does not require that an animal perform the response to be eliminated, it could not produce an increase in the inhibitory factors that oppose reaction potential, and consequently could have no effect upon response strength. But while it is clear that Hull's response-induced inhibition theory of extinction requires the performance of the response to be eliminated, and is thus not adequate to account for the effects of latent extinction, Hull's general behavior theory makes no such requirement. In a series of early papers (15, 16, 17, 18, 19, 20), Hull explored the possibility of applying a number of explanatory constructs to complex behavioral relationships in which reasoning, insight, knowledge, and purpose appeared to be operative. One of these constructs—the fractional anticipatory goal response—was used subsequently by others to mediate a wide range of behavior phenomena. For example, Kendler's analysis of switching behavior (25, 26, 27) and Osgood's mediation hypothesis (35) both rely heavily on the explanatory capacity of the fractional anticipatory response mechanism. Following the leads contained in Hull's early theoretical article, Spence (40, 41, 43) has provided a somewhat more rigorous analysis of the role played by this mechanism in the acquisition of a T-maze choice response, and has used this analysis to integrate latent-learning data and data concerned with the acquisition

of opposed spatial responses under the simultaneous presence of hunger and thirst. It is the opinion of the present author that the fractional anticipatory goal response possesses the deductive potential necessary to mediate also an interpretation of latent extinction data that is consistent with an S-R reinforcement approach.

Spence has suggested (40) that, as a function of differential reinforcement during training, the stimulus traces of the cues in the positive goal box and those in the alley leading to it become (classically) conditioned to the goal response (e.g., eating or drinking). Through generalization, cues at the entrance to the correct alley acquire the capacity to evoke that fractional component of the goal response ( $r_g$ ) not in conflict with the overt acts of the behavior sequence. This component produces a characteristic proprioceptive stimulus ( $s_g$ ) which in turn becomes conditioned to entering and continuing locomotion in the alley leading to the positive goal box. Through the acquisition of this associative connection,  $s_g$  becomes a component of the habit strength determining the instrumental response. It has also been suggested that "through the intensity of its trace which . . . increases with the increased strength of  $r_g$  as the latter becomes more strongly conditioned,  $s_g$  determines the non-associative factor  $K$ " (incentive motivation) (40, p. 273). Spence implies that the magnitude of this nonassociative component is an increasing monotonic function of the intensity of  $s_g$ . Since the latter, in turn, is assumed to covary positively with the strength of  $r_g$ , such variables as the amount of reinforcement and the delay of reinforcement (which presumably affect the strength of  $r_g$ ) play a role in determining the value of  $K$ . If it is recalled that

$$S^E R = f(M \times S^H R),$$



where

$$M = f(K + V + D),$$

it can be seen that  $r_g$  and its characteristic proprioceptive cue determine the excitatory strength of the correct response by contributing a component to  $S^H R$  and also by determining the value of  $K$ .<sup>3</sup>

But while the fractional anticipatory goal response has been considered by some students of learning to be a construct that possesses a great deal of explanatory fertility, it has been considered by others as an *ad hoc* device designed to rescue the S-R theorist from embarrassing experimental data. Koch, for example, maintains that:

Among the potentially most fruitful items in Hull's bequest to the future are the many ingenious "peripheral" mechanisms which were elaborated in the first instance, in his earlier theoretical articles. . . . It is likely that concepts of this order point to factors which must be ultimately taken into account by any theory of behavior . . . (28, pp. 164-165).

On the other hand, Meehl and MacCorquodale write that ". . .  $r_g$  at present . . . is as readily available a *deus ex machina* for nonexpectancy theorists as the concepts of 'attention,' 'emphasis,' or 'perceptual threshold' are for Tolman" (31, p. 232). Since  $r_g$  appears to enjoy a somewhat ambiguous epistemological position, it would be well to clarify its meaning before attempting to provide an interpretation of the effects of latent extinction.

The use of a theoretical construct denoting a nonobservable response process carries with it the responsibility of pro-

viding a set of symptom relations by which the construct in question can be coordinated with observable events. A number of such reductions to data language is essential if the construct is to acquire univocal intersubjective reference, and if it is to be integrated with other constituents of the theoretical system in which it functions. One way in which a set of reduction symptoms can be provided for  $r_g$  is to conceive its relation to certain antecedent manipulable variables to be the same as that of any other molar response with which contemporary behavior theory has been concerned (e.g., bar pressing, spatial responses, salivation, etc.). This conceptualization implies that such variables as, for example, number of reinforced trials, quality and magnitude of the reward object, and hours of deprivation will exert the same effect on the strength of  $r_g$  as they have been found to exert on the strength of these molar responses. Once the relations postulated to hold between  $r_g$  and certain experimental parameters are clearly set forth it becomes possible to specify the effect upon  $r_g$  of variations in the value of any one of these parameters. However, it should be emphasized that, since  $r_g$  is an inferred response and as such does not possess a directly observable element, evidence relevant to variations in the strength of  $r_g$  can be obtained only by examining the overt behavior of the organism under certain sets of experimental conditions.

### III

*The Type I situation.* Consider the situation in which an animal is required to turn right in a T maze to obtain a reward object relevant to an existent need state. When the excitatory tendency of the cues in the right alley to evoke a right turn becomes greater than that of the cues in the left alley to evoke a left turn, the animal will respond

<sup>3</sup> The additive relationship between  $V$ ,  $K$ , and  $D$  presented here is consistent with Spence's latest formulation (42). Following Hull, Spence had earlier postulated a multiplicative relationship between these factors, the implication being, of course, that if any one of them is reduced to zero the excitatory strength of the instrumental response would also be reduced to zero.

above chance expectancy in the direction of the goal location. An important factor involved in the growth of the correct excitatory potential ( $S^E R_a$ ) is the increase in habit strength ( $S^H R$ ) resulting from each reinforcement. However, as the animal is repeatedly reinforced during the course of training, not only does  $S^E R_a$  approach its asymptote but, in addition, the overt consummatory response appropriate to the reward object (i.e., eating or drinking) becomes more strongly conditioned to the stimulus traces of the cues in the goal location. On the basis of the latter it appears reasonable to assume that when latent extinction is administered following the termination of training, that component of the consummatory response which can occur in the absence of the reward object (i.e.,  $r_g$ ) will be evoked by the cues to which its molar counterpart was previously conditioned. Since these cues are constantly present during latent extinction, it would be expected that  $r_g$  will be strongly and repeatedly evoked; each latent extinction "trial" providing the occasions for its *free repetition*.<sup>4</sup> Now if  $r_g$  is conceived as a response process that obeys the same behavioral laws as most molar systems, then it would follow that these successive non-reinforced emissions of  $r_g$  will produce a sharp reduction in its response strength with respect to cues in the goal location. If these cues are similar to those at the choice point, and if delay of reward during training was at a minimum, the extinction effects of  $r_g$  will generalize readily to antedating segments of the behavior sequence, so that at the start of the test trials  $r_g$  either will not be elicited at the choice point or will be

elicited with greatly reduced excitatory potential. If latent extinction is effective in reducing the strength of  $r_g$  to the extent that it is subsequently not available at the choice point, then  $s_g$ , of course, will also not be available, and the associative connection established during training between  $s_g$  and the instrumental response will no longer contribute to the habit strength of that response. As a consequence of habit withdrawal of this component of habit strength, the organism will manifest a decreased tendency to enter the correct alley during the test trials.

But assume that  $r_g$  was not completely extinguished during the latent extinction period, and that at the start of the test trials the cues at the choice point still possessed some excitatory tendency with respect to  $r_g$ . Even if latent extinction only reduced this tendency, a decrease in the strength of the established instrumental response would nevertheless be expected. It will be recalled that the incentive motivational factor  $K$  was postulated to be a nonassociative component of  $S^E R_a$ . Since the relationship between  $K$  and  $s_g$  is assumed to be monotonically increasing and, furthermore, since the intensity of  $s_g$  is a direct function of the  $r_g$  evoked, it would follow that any decrease in the strength of  $r_g$  will reduce the value of  $K$ . This reduction of  $K$  will, in turn, reduce the excitatory potential mediating the correct approach response, as a result of which the animal will show a decreased tendency to enter and continue locomotion in the alley leading to the previously positive goal box.

*The Type II situation.* The experimental procedures that we have labeled Type I and Type II are similar in that they both involve introducing an animal into an empty goal location after it has acquired a response instrumental in securing a reward object previously contained in that location. The difference

<sup>4</sup> Failure to conceive of latent extinction as providing the occasions for the *free repetition* of  $r_g$  led Gleitman, Nachmias, and Neisser (14) to dismiss the possibility of employing  $r_g$  to interpret the results of those experiments in which a Type I procedure was used.



between these procedures arises with respect to the responses with which they deal following the latent extinction period. It will be recalled that while the Type I procedure is concerned with the effect of latent extinction on the strength of an established response, the Type II procedure is concerned with the capacity of goal-box stimuli that have acquired secondary reward value to mediate the learning of a *new response* after latent extinction has been administered. It is to the latter procedure that we now turn our attention.

Consider the experimental situation in which an animal is trained to traverse a straight alley in order to obtain a reward object relevant to an existent need state. A differential reinforcement technique is employed by which, on a given number of trials, the animal runs to a white goal box that contains the reward object, while on the remaining trials he runs to an empty black goal box. By the principle of secondary reinforcement the cues in the white goal box will acquire reward value as a function of their association with the reinforcement situation. Hull (21, 22) suggested the fractional anticipatory goal response and its proprioceptive cue as providing the mechanism underlying the acquisition of secondary reward properties. The role that  $r_g$  plays in this respect develops by virtue of its intimate association with the reinforcement situation, this association establishing  $s_g$  as a secondary reinforcer. In order for any neutral stimulus to acquire reward value it must acquire the capacity to elicit  $r_g$  either by being conditioned to  $r_g$  directly, as in the case of stimuli that are spatially and temporally contiguous with the receipt of the reward object, or through generalization, as in the case of stimulus events that are similar to those actually conditioned. Hull further suggested that a stimulus that has acquired reward value will retain that value as

long as it retains the power to evoke  $r_g$ . It would appear reasonable on this basis to assume the acquired reward capacity of a stimulus to be some positive function of the strength of the  $r_g$  to which it is conditioned, and for any subsequent reduction in this strength to decrease the reinforcing value of the stimulus.

Now with respect to the experimental situation being considered, assume that following training the positive and negative goal boxes are placed on opposite arms of a T maze, and that the animal is given a number of free trials in the *absence of the reward object*. When he first responds correctly and enters the positive goal box,  $r_g$  will be evoked by those stimuli to which it was conditioned during the course of training. As a result of the evocation of  $r_g$ , and the presence consequently of  $s_g$  with its reinforcing capacity, there will occur an increment in the excitatory potential of the correct response. As long as  $r_g$  continues to be evoked in the positive goal box, the correct response will be reinforced and will increase in strength relative to the incorrect response. But the excitatory potential of the correct response will not continue to increase unabated until some asymptotic value is reached. Since the reward object is not present during the maze series, the emission of  $r_g$  in the positive goal box following each correct choice will reduce its response strength, as a result of which the intensity of  $s_g$  and hence the magnitude of available secondary reinforcement will decrease progressively. When the stimuli in the positive goal box no longer elicit a threshold value of  $r_g$ , the excitatory potential of the correct response will begin to decrease, so that after several trials the animal will cease to respond correctly in excess of chance.

Now assume that latent extinction is administered at some time between the termination of straight-alley training and the beginning of the maze trials. For



reasons presented in connection with the Type I situation, placing an animal in a previously baited goal box provides the occasion for the free nonreinforced repetitions of  $r_g$ , each such repetition reducing its response strength. Because of the role that  $s_g$  plays as a reinforcing agent, it would be expected that, as the strength of  $r_g$  decreases during the course of latent extinction, the intensity of  $s_g$ , and hence the secondary reward value of goal-box stimuli, will also decrease. Depending on the duration of latent extinction and the presence of other conditions to be discussed below,  $r_g$  either will not be elicited in the positive goal box at the termination of the latent extinction period or it will be elicited with greatly reduced excitatory potential. In the event that  $r_g$  is completely extinguished with respect to goal-box stimuli, there will be no secondary reinforcement available to mediate subsequently the learning of a choice response in the maze. On the other hand, if the excitatory strength of  $r_g$  had not been eliminated but had been reduced to a near threshold value, then  $r_g$  will become completely extinguished *early* in the maze series. In any case, animals that have been given latent extinction would not be expected to show a strong tendency to choose the arm of the maze leading to the positive goal box, as compared with control animals that had not undergone the latent extinction experience.

By employing the fractional anticipatory response mechanism in a manner consistent with the requirements of the neobehavioristic system in which it was developed, we have been able to provide an interpretation of the effects of latent extinction in both the Type I and Type II situations. Within the theoretical framework employed it was necessary only to assume that  $r_g$  is repeatedly evoked during latent extinction, and that, as a consequence of this

repeated evocation,  $r_g$  either becomes completely extinguished with respect to cues in the goal location or that a sharp reduction in its response strength occurs.

#### IV

The value of a theoretical construct should be measured not only in terms of how well propositions involving that construct explain available data but also in terms of the extent to which these propositions generate novel (i.e., not previously formulated) functional relationships. It must be determined, therefore, whether the fractional anticipatory response mechanism as employed in the present analysis is capable of mediating predictions concerning the influence on behavior of certain selected variations in both the Type I and Type II situations. Unless the present analysis possesses a prospective reference, propositions involving  $r_g$  have no value other than as vehicles in providing an economical description of already established empirical relationships. While such descriptions play an important role in science, the constructs involved function simply as analytic devices whose systematic significance is limited to furnishing summary statements of observed facts (5).

On the basis of our analysis of the effect of latent extinction on the strength of  $r_g$ , we can specify the following general hypothesis: *Latent extinction will reduce the excitatory strength of an instrumental response to the extent that it reduces the excitatory strength of  $r_g$ .* Following from this is the implication that if a variable is introduced, during latent extinction, that increases the degree to which  $r_g$  is weakened, a greater decrement in subsequent molar performance will occur than if latent extinction is administered in the absence of this variable. It will be recalled that  $r_g$  was conceived as a hypothetical response process whose relation to observable

antecedent events was assumed to follow the same behavioral laws as molar responses like bar pressing and salivation. This implies that all those experimental operations that have been demonstrated to reduce the strength of a molar response should affect similarly the strength of  $r_g$ . Since empirical evidence appears to indicate that reduction in the strength of these molar responses occurs as a monotonic increasing function of the number of nonreinforced evocations, it would follow that the more frequently  $r_g$  is evoked during latent extinction the greater will be the reduction in its response strength. On this basis we would be led to expect latent extinction to produce a pronounced decrement in subsequent molar performance under those experimental conditions that produce a rapid rate of emission of  $r_g$  during the latent extinction period. We turn now to consider some of these conditions.

*Variations in the strength of the relevant drive.* The results of several studies have indicated that the higher the level of a relevant (rewarded) drive during a series of extinction test trials the greater the rate of response evocation. For example, Jenkins and Daugherty (24) have shown that following training an increase in drive strength will result in the more frequent emission of a nonreinforced pecking response, while Skinner (39) has presented evidence confirming this relationship with respect to bar pressing. On the assumption that  $r_g$  obeys the same behavioral laws as these responses, we would expect a higher drive level during latent extinction to result in a greater increase in the number of nonreinforced evocations of  $r_g$ .<sup>5</sup> On this basis the following hypothesis can be specified with reference to the Type I

situation: If after being trained to a position habit in a T maze under identical motivational conditions, one group of animals is given latent extinction under a high drive and another group under a low drive, the high-drive group will show a greater decrease in the tendency to choose the previously baited goal box when both groups are again run under the motivational level that prevailed during acquisition. A recent experiment by Pliskoff (36) is relevant to the present hypothesis. Animals under 8 hr. of food deprivation were trained in the performance of a T-maze turning response, after which latent extinction was administered under either 0, 8, or 22 hr. of hunger. When the extinction test trials were subsequently run under acquisition conditions it was found, in accord with expectation, that on the first test trial the 22-hr. group chose the previously correct side of the maze less often than the 0-hr. group. This result suggests a relationship between drive level during latent extinction and reduction in the strength of a previously established response.

The assumption concerning the relation between drive strength and rate of emission of  $r_g$  is also relevant to the Type II situation. It will be recalled that the acquired reward capacity of a neutral stimulus was assumed to be some positive function of the excitatory strength of the  $r_g$  to which that stimulus was conditioned. This implies that if increasing the drive level during latent extinction serves to increase the number of nonreinforced emissions of  $r_g$  (and thereby to decrease the strength of the latter with respect to goal location stimuli), then there should be a more pronounced reduction in secondary reward value the more intense the drive level under which latent extinction is administered. Evidence supporting such a hypothesis was reported recently by Moltz and Maddi (33). They employed

<sup>5</sup> The relation between drive strength and response evocation also follows readily from Hull's 1943 (21) and 1952 (22) postulate sets.



a differential reinforcement technique during training which consisted of running animals in a straight alley to a goal box of one color on rewarded trials, and to a goal box of a different color on nonrewarded trials. While all Ss were under the same hunger drive schedule during training (22 hr.), they were under either a 0-, 22-, or 44-hr. drive at the beginning of the latent extinction period. The test trials were subsequently run in a T maze, with all Ss again on the same hunger deprivation schedule imposed during training. It was found that the mean number of responses (in 15 free choices) to the side of the maze containing the positive goal box was 9.30, 7.10, and 6.30 for the 0-, 22-, and 44-hr. groups, respectively.<sup>6</sup> These results are in accord with theoretical expectation, since they indicate a relationship between motivational level during latent extinction and the capacity of goal location stimuli to mediate subsequently the learning of a new response.

*Variations in the strength of an irrelevant drive.* Hull has proposed that the drive strength available at any given amount is a function not only of the need for which the goal object in the situation is appropriate but also of all irrelevant (nonrewarded) drives operative at the time. It is the total need state, consisting of both relevant and irrelevant drives, that is assumed to combine with  $S^H R$  (in addition to the nonassociative factors  $V$  and  $K$ ) to determine excitatory potential. In *Principles of Behavior* (21) Hull stated as a corollary that, with the number of reinforcements held constant, response strength will increase as some positive

function of the magnitude of an irrelevant drive.

Webb (46) reported the results of a study designed to determine the role of an irrelevant drive (thirst) in activating a habit structure when the drive under which the habit was originally developed (hunger) was satiated. He found the strength of a panel-pushing response to be approximately an increasing linear function of the intensity of the irrelevant drive. Brandauer (6) subsequently confirmed this relationship, using a different operant response. Since we have assumed that  $r_p$  follows essentially the same behavioral laws as most molar responses, we appear to be in the somewhat uncomfortable position of having to propose that  $r_p$  appropriate to a relevant hunger drive will increase in strength in approximately linear fashion with increases in the magnitude of an irrelevant thirst drive. However, the implications of a theory or of a postulate are determined not only by the relations or laws specified but also by the conditions operative in the experimental situation to which these relations or laws are applied. As a case in point, consider the situation into which there is introduced an irrelevant need that provides a drive stimulus which elicits response *incompatible* with a previously established criterion response. Under such conditions it would not appear reasonable to expect the strength of the criterion response to increase as the irrelevant need becomes more intense. Indeed, only when conflicting habit tendencies are not activated would a criterion response, whether it be "molar" or "molecular," be expected to increase with increases in the intensity of an irrelevant drive stimulus.

The results of several studies (1, 2, 7, 13) in the area of secondary motivation provide evidence in support of this contention. For example, Amsel and Maltzman (3) have shown that, when

<sup>6</sup> The  $F$  ratio for the difference between these drive levels was significant well below .01. No significant difference in maze performance was found between control groups which were not given latent extinction but which were made differentially hungry during the latent extinction period.



anxiety is added to an existing motivational complex, the strength of a previously established consummatory (drinking) response will increase if the experimental situation is arranged so that responses conditioned to the irrelevant anxiety-drive stimulus (e.g., crouching, face washing, escape responses, etc.) cannot compete with the drinking response. On the other hand, if the situation is arranged so that response competition is permitted, the rate of drinking will be sharply reduced. These results, in addition to those provided by Kendler (26) concerning the "incompatibility" of fractional anticipatory eating and drinking responses, make it appear reasonable to assume that anticipatory drinking responses ( $r_d$ ), when instigated by an irrelevant thirst drive, will compete with anticipatory eating responses ( $r_e$ ), and consequently will decrease the rate of emission of the latter.<sup>7</sup> Following from this is the implication that after training under moderately intense hunger, the stronger the thirst drive that is subsequently introduced (thereby increasing the strength of  $r_d$ ) the greater the reduction in the rate of evocation of  $r_e$ .

This analysis generates an important deduction with respect to both the Type I and Type II situations. Assume that an animal learns to approach a distinctive goal location under moderately in-

<sup>7</sup> On the assumption that anticipatory drinking responses are in some manner "incompatible" with anticipatory eating responses, Kendler has attempted to explain what he has called the switching phenomenon, i.e., "... the tendency of rats after learning to go to one end of a T maze to obtain dry food to switch their spatial responses when made thirsty" (26, p. 179). He tested several deductions based on this assumption, one of which was that the conflict between  $r_e$  and  $r_d$  would be intensified (and hence the amount of switching increased) when  $S$  is shifted to a strong as compared with a weak thirst drive. The results obtained were consistent with this deduction.

tense hunger, and that as a function of this training the cues in the goal location acquire a strong excitatory tendency with respect to  $r_e$ . If latent extinction is administered subsequently under both hunger and thirst, the presence of incompatible anticipatory drinking responses should reduce the number of nonreinforced evocations of  $r_e$ ; this reduction being greater the more intense the thirst drive. On this basis, and on the basis of the previously formulated assumption concerning relation between rate of emission of  $r_e$  and the influence of latent extinction on molar performance, we can state the following general hypothesis: If an organism is trained to acquire a response under moderately strong hunger but with ad libitum consumption of water, the effectiveness of latent extinction in reducing the strength of this response will decrease as the intensity of thirst during latent extinction increases. More specifically, in the Type I situation we would expect that if, after learning a position habit under 22 hr. of hunger, one group of animals is given latent extinction under the same hunger drive but in conjunction with 3 hr. of thirst, and another group in conjunction with 22 hr. of thirst, the latter will yield significantly more responses to extinction when the motivational complex of both groups is subsequently made identical with that which prevailed during acquisition.

Since the relation between intensity of an irrelevant thirst drive and the rate of emission of  $r_e$  would be expected to obtain in a Type II situation as well, we may also state the following hypothesis: After an approach response to food is established under moderately intense hunger, a greater reduction in the acquired reward value of goal location stimuli will occur if, in conjunction with the prevailing hunger drive, latent extinction is administered under 3 hr. of thirst as compared with 22 hr. of thirst.

Unfortunately there is no empirical evidence relevant either to this hypothesis or to the hypothesis formulated with reference to the Type I situation. Several experiments are, however, being conducted in the writer's laboratory in the attempt to determine the role of drive interaction during latent extinction.<sup>8</sup>

*Reward immediately prior to latent extinction.* In the presence of the appropriate drive state, a reward object with which an animal has had previous commerce will serve typically as an "unconditioned" stimulus for the consummatory response, in the sense that the latter will be regularly elicited when the reward object is presented. Since  $r_g$  has been conceived as a hypothetical process representing a fractional component of the consummatory response, it appears reasonable to assume that  $r_g$  possesses this functional characteristic in common with its molar counterpart. On this basis we would expect that, with the primary drive state remaining relatively constant, the frequency with which  $r_g$  is evoked will increase during a period following consumption of an appropriate goal object. However, if a large amount of reward is consumed, a reduction in the primary drive state would occur which, in turn, might offset or prevent entirely an increase in the rate of emission of  $r_g$ . On the other hand, a small amount of reward, while not appreciably modifying the drive state, might not be sufficient to effect this increase. Presumably there is an optimal amount of reward that would produce the desired result. With respect to food as the reward object, experimental evidence reported by Maltzman (30) indicates this amount to be within the range of 1000 to 2000 mg. when the animal is under a hunger drive

of approximately 22 hr.<sup>9</sup> On the basis of Maltzman's results, we would be led to expect that if, after training under hunger motivation, an animal is fed about 1500 mg. of food 20 sec. prior to latent extinction, an increase in the rate of emission of  $r_g$  will occur during the latent extinction period. This increase in rate of emission will, according to the present analysis, produce a marked reduction in the strength of  $r_g$ , as a consequence of which latent extinction should be more effective in reducing subsequent molar performance. Thus in the Type I situation it would be expected that if latent extinction is preceded by the ingestion of a reward object appropriate to the prevailing drive state, a greater decrement in the strength of the established instrumental response will occur, while in the case of the Type II situation a greater decrement in the acquired reward value of goal location stimuli would be expected.

The hypotheses that have been presented in the present paper have been concerned with the relation between certain experimental variables introduced during latent extinction and the rate of emission of  $r_g$ . Further experimental work and theoretical analysis might also be directed toward determining whether variations in molar performance following latent extinction can be ascribed to other parameters of  $r_g$  function. In addition, the effect of latent extinction might also be investigated with regard to the influence of such training parameters as the quality and magnitude

<sup>9</sup> Maltzman found that hungry rats fed about 1200 mg. of food approximately 20 sec. before being run in a straight alley tended to run significantly faster than a control group that was not pre-fed. Several other investigators (4, 8, 34) have also reported that small amounts of preliminary reward will facilitate the performance of an instrumental response. Maltzman has indicated how these results can be interpreted in terms of the fractional anticipatory response mechanism.

<sup>8</sup> These studies are being carried out in collaboration with Miss Nina Tokarow and Mr. Leonard Rosenblum.



of the reward object, number of reinforced trials, and the nature of contextual stimulus conditions.

## V

The present paper represents an attempt to provide an interpretation of latent extinction that is consistent with an S-R reinforcement approach. The effects of latent extinction on the strength of an established instrumental response and on the secondary reward value of goal location stimuli were analyzed. It was found that propositions involving the fractional anticipatory goal response in conjunction with other propositions of neobehavioristic theory are adequate to account for these effects. Several hypotheses were suggested concerning the influence of certain variables on the rate of emission of the anticipatory goal response during the latent extinction period. Evidence relevant to these hypotheses was considered.

## REFERENCES

1. AMSEL, A. The effect upon level of consummatory response of the addition of anxiety to a motivational complex. *J. exp. Psychol.*, 1950, 40, 709-715.
2. AMSEL, A., & COLE, K. F. Generalization of fear-motivated interference with water intake. *J. exp. Psychol.*, 1953, 46, 243-247.
3. AMSEL, A., & MALTZMAN, I. The effect upon generalized drive strength of emotionality as inferred from the level of consummatory response. *J. exp. Psychol.*, 1950, 40, 563-569.
4. ANDERSON, E. E. The externalization of drive. IV. The effect of prefeeding on the maze performance of hungry rats. *J. comp. physiol. Psychol.*, 1941, 31, 349-352.
5. BECK, L. W. Constructions and inferred entities. In H. Feigl and May Brodbeck (Eds.), *Readings in the philosophy of science*. New York: Appleton-Century-Crofts, 1953.
6. BRANDAUER, C. M. A confirmation of Webb's data concerning the action of irrelevant drives. *J. exp. Psychol.*, 1953, 45, 150-152.
7. BROWN, J. S., KALISH, H. I., & FARBER, I. E. Conditioned fear as revealed by magnitude of startle response to an auditory stimulus. *J. exp. Psychol.*, 1951, 41, 317-328.
8. BRUCE, R. H. An experimental investigation of the thirst drive in rats with especial reference to the goal-gradient hypothesis. *J. gen. Psychol.*, 1937, 17, 49-60.
9. BUGELSKI, B. R., COYER, R. A., & ROGERS, W. A. A criticism of pre-acquisition and pre-extinction of expectancies. *J. exp. Psychol.*, 1952, 44, 27-30.
10. COATE, W. B. Weakening of conditioned bar-pressing by prior extinction of its subsequent discriminated operant. *J. comp. physiol. Psychol.*, 1956, 49, 135-138.
11. DEESE, J. The extinction of a discrimination without performance of a choice response. *J. comp. physiol. Psychol.*, 1951, 44, 362-366.
12. DEESE, J. *The psychology of learning*. New York: McGraw-Hill, 1952.
13. ESTES, W. K., & SKINNER, B. F. Some quantitative properties of anxiety. *J. exp. Psychol.*, 1941, 29, 390-400.
14. GLEITMAN, H., NACHMIAS, J., & NEISSER, U. The S-R reinforcement theory of extinction. *Psychol. Rev.*, 1954, 61, 23-33.
15. HULL, C. L. Knowledge and purpose as habit mechanisms. *Psychol. Rev.*, 1930, 37, 511-525.
16. HULL, C. L. Goal attraction and directing ideas conceived as habit phenomena. *Psychol. Rev.*, 1931, 38, 487-506.
17. HULL, C. L. The concept of the habit-family hierarchy and maze learning. *Psychol. Rev.*, 1934, 41, 33-54, 134-152.
18. HULL, C. L. The mechanism of the assembly of behavior segments in novel combinations suitable for problem solution. *Psychol. Rev.*, 1935, 42, 219-245.
19. HULL, C. L. Mind, mechanism, and adaptive behavior. *Psychol. Rev.*, 1937, 44, 1-32.
20. HULL, C. L. Fractional antedating goal reactions as pure stimulus acts. In *Psychological Memoranda, 1940-1944*. Bound mimeographed material on file in the libraries of the Univer. of Iowa, Univer. of North Carolina, and Yale Univer., Oct. 24, 1941.
21. HULL, C. L. *Principles of behavior*. New York: Appleton-Century-Crofts, 1943.
22. HULL, C. L. *A behavior system*. New Haven: Yale Univer. Press, 1952.



23. HURWITZ, H. M. B. Response elimination without performance. *Quart. J. exp. Psychol.*, 1955, 7, 1-7.
24. JENKINS, W. O., & DAUGHERTY, GEORGETTE. Drive and the asymptote of extinction. *J. comp. physiol. Psychol.*, 1951, 44, 372-377.
25. KENDLER, H. H., & LEVINE, S. Studies of the effect of change of drive: I. From hunger to thirst in a T-maze. *J. exp. Psychol.*, 1951, 41, 429-436.
26. KENDLER, H. H., LEVINE, S., ALTCHER, E., & PETERS, H. Studies of the effect of change of drive: II. From hunger to different intensities of a thirst drive in a T-maze. *J. exp. Psychol.*, 1952, 44, 1-3.
27. KENDLER, H. H., KARASIK, A. D., & SCHRIER, A. M. Studies of the effect of change of drive: III. Amounts of switching produced by shifting from thirst to hunger and from hunger to thirst. *J. exp. Psychol.*, 1954, 47, 179-182.
28. KOCH, S. Clark L. Hull. In W. K. Estes, S. Koch, K. MacCorquodale, P. E. Meehl, C. G. Mueller, Jr., W. N. Schoenfeld, and W. S. Verplanck, *Modern learning theory*. New York: Appleton-Century-Crofts, 1954. Pp. 1-176.
29. LEVY, N. Latent extinction of a jumping response. *Amer. Psychologist*, 1955, 10, 416. (Abstract)
30. MALTZMAN, I. The process need. *Psychol. Rev.*, 1952, 59, 40-48.
31. MEEHL, P. E., & MACCORQUODALE, K. Some methodological comments concerning expectancy theory. *Psychol. Rev.*, 1951, 58, 230-233.
32. MOLTZ, H. Latent extinction and the reduction of secondary reward value. *J. exp. Psychol.*, 1955, 49, 395-400.
33. MOLTZ, H., & MADDI, S. R. Reduction of secondary reward value as a function of drive strength during latent extinction. *J. exp. Psychol.*, 1956, 52, 71-76.
34. MORGAN, C. T., & FIELDS, P. E. The effect of variable preliminary feeding upon the rat's speed of locomotion. *J. comp. Psychol.*, 1938, 26, 331-348.
35. OSGOOD, C. E. *Method and theory in experimental psychology*. New York: Oxford Univer. Press, 1953.
36. PLISKOFF, S. S. Response elimination as a function of generalization, motivation and number of non-response extinction trials. Unpublished doctor's dissertation, New York Univer., 1955.
37. SCHARLOCK, D. P. The effects of a pre-extinction procedure on the extinction of place and response performance in a T maze. *J. exp. Psychol.*, 1954, 48, 31-36.
38. SEWARD, J. P., & LEVY, N. Sign learning as a factor in extinction. *J. exp. Psychol.*, 1949, 39, 660-668.
39. SKINNER, B. F. Are theories of learning necessary? *Psychol. Rev.*, 1950, 47, 193-216.
40. SPENCE, K. W. Theoretical interpretations of learning. In C. P. Stone (Ed.), *Comparative psychology*. New York: Prentice-Hall, 1951. Pp. 239-291.
41. SPENCE, K. W. Theoretical interpretations of learning. In S. S. Stevens (Ed.), *Handbook of experimental psychology*. New York: Wiley, 1951. Pp. 690-729.
42. SPENCE, K. W. Current interpretations of learning data and some recent developments in stimulus-response theory. In *Learning theory, personality theory, and clinical research. The Kentucky Symposium*. New York: Wiley, 1954.
43. SPENCE, K. W., BERGMANN, G., & LIPPITT, R. A study of simple learning under irrelevant motivational-reward conditions. *J. exp. Psychol.*, 1950, 40, 539-551.
44. TOLMAN, E. C. *Purposive behavior in animals and men*. Berkeley: Univer. of California Press, 1932.
45. TOLMAN, E. C. There is more than one kind of learning. *Psychol. Rev.*, 1949, 56, 144-155.
46. WEBB, W. B. The motivational aspect of an irrelevant drive in the behavior of the white rat. *J. exp. Psychol.*, 1949, 39, 1-14.

(Received August 29, 1956)

## THE CELL ASSEMBLY: MARK II<sup>1</sup>

P. M. MILNER

McGill University

The purpose of this paper, as its title implies, is to suggest a revision of the neural construct that Hebb (5) has called a "cell assembly." The cell assembly plays a fundamental role in Hebb's behavior system, especially in his admirable treatment of perceptual learning, but the proposed changes are not intended to affect these psychological implications of the construct; they are aimed, rather, at making the underlying neural processes more explicit. As Hebb has mentioned (8), he tried to limit himself to using only known, or highly probable, properties of neurons in setting up his model; but it is clear that these soon proved to be inadequate, and he then had to introduce *ad hoc* mechanisms to explain how such artificially simplified elements could combine to produce the observed molar behavior. We can now see that from many standpoints the theory would have been better if the additional postulates had been introduced at the neural level, where they could have been more easily evaluated and tested; but on the other hand, such a theory might never have been given a hearing.

There is little doubt that flaws exist in the arguments Hebb uses to derive the molar properties of his construct. As will be shown below, if we try to operate with only those properties and connections of individual neurons that

Hebb considered respectable, the behavior of the resulting network is incompatible with that required for the cell assembly. Some of the inconsistencies can be eliminated simply by taking into account neurophysiological discoveries made during the last ten years, but in order to derive a system having anything like the characteristics required for the cell assembly it will be necessary also to invoke a number of purely hypothetical neural processes.

As far as can be ascertained from *The Organization of Behavior* (5), only four factors are considered to determine whether a cortical neuron will fire: (a) the number of impulses bombarding the neuron from all sources during the few milliseconds in which temporal summation is assumed to take place; (b) the "strength" of the synapses concerned (the strength of a synapse may increase beyond its native value in accordance with the provisions of a learning postulate); (c) whether the neuron is refractory; and (d) by implication at least, the effect of neural fatigue. In the network that Hebb postulated as a model of the cortex, neurons with the above-mentioned properties are randomly interconnected, each neuron having efferent and afferent connections with many others. Some of the neurons can be fired by sensory input.

Such a network is somewhat analogous to the nuclear fission reactor. In that machine a neutron may either induce a fission, in which case it releases several new neutrons, or it may be absorbed or lost without releasing any other neutrons. If the neutron flux in the machine is to remain constant, all the neutrons released at a fission must

<sup>1</sup> The writer wishes to acknowledge the influence on this paper of many helpful arguments with his colleagues at McGill, especially those with Drs. James Olds and Seth Sharpless, and with Professor D. O. Hebb. The paper was written while the author was receiving support from the Foundations' Fund for Research in Psychiatry, and from the Ford Foundation.

be lost except one. For stable operation, in other words, the multiplication factor (the average number of fissions produced by the neutrons released at a previous fission) must be one. Likewise, in the cortical network, a fired neuron may fire several others, or it may send impulses to refractory, fatigued, or inadequately facilitated neurons and produce no further firing. If the average neuron fires one other neuron, the total activity in the network will remain constant.

The analogy is not quite complete, however; in the fission reactor the multiplication factor is slightly reduced by an increase in the neutron flux, but in the cortical network just the opposite is true, the multiplication factor increasing with total neural activity. This results from the effect of summation at synapses. When a neuron is firing by itself, the probability of its firing another is small; but when many neurons are firing at the same time, they provide a facilitatory background for each other, and one of them may then be able to fire several neurons to which it sends efferents. Thus, the number of neurons fired by an average neuron will increase in proportion to the total number of neurons firing, a state of affairs which precludes stable operation. In the nuclear reactor, if the multiplication factor should fluctuate to a value of 1.001, a few more neutrons will be released than are lost, and the total flux will increase; at this higher flux the multiplication factor will be reduced again to one. If the same fluctuation should occur in the cortical network, a few more impulses will circulate, the general level of facilitation will rise, and the multiplication factor will increase still further. This increase will only be halted when all the neurons in the system are firing as fast as their refractory periods will allow. Unlike the nuclear pile, therefore, the cortical network has

only two levels of stable operation, "full on" and "full off." Neither of these states fulfills the requirements of Hebb's cell assembly. It is hardly necessary to add that the phase sequence, which is essentially a chain of associated cell assemblies representing a "train of ideas," bears as tenuous a relationship to the basic neural axioms as does the cell assembly itself.

The fact that Hebb's constructs are not rigorously derived is not as serious for the over-all theory as it might seem, because many of the more molar physiological principles that Hebb introduces are valid and important for the understanding of psychological phenomena, irrespective of the exact neural mechanisms involved. Nevertheless, the possibility of bridging the gap between the physiological and the psychological levels with a more substantial structure presents a fascinating challenge, especially in view of the improvements made to the supports on the physiological side since *The Organization of Behavior* was written.

#### *A New Physiological Interpretation of the Cell Assembly*

In what follows it will be recognized that we are adopting Hebb's descriptions of the cell assembly and the phase sequence as far as their psychological properties are concerned, but instead of trying to reconcile these to a minimum of oversimplified neural data, further postulates will be introduced in an attempt to provide an explicit relationship between the constructs and their elements.

The first additional postulate concerns inhibition. At the time when Hebb was developing his theory, many physiologists were strongly opposed to the idea of neural inhibition, largely because it was difficult to fit into the electrical theory of synaptic transmission. There was a feeling that all inhibitory phe-



nomena could be explained by lack of facilitation, or by invoking the principle of refractoriness, and Hebb seems to have accepted this view. Since then, the work of Eccles (3) and others has provided a better understanding of synaptic transmission, and inhibitory transmission is now generally accepted.

The processes underlying learning have still not been revealed by the neurophysiologists, and to fill this lacuna there seems to be nothing against retaining the postulate suggested by Hebb: "When an axon of cell A is near enough to excite a cell B and repeatedly or persistently takes part in firing it, some growth process or metabolic change takes place in one or both cells such that A's efficiency, as one of the cells firing B, is increased" (5, p. 62).

Another assumption that we shall make here is that the cell assembly involves mainly cortical neurons. It is to be understood, however, that connections between cortical cells may often take place via cells in the thalamus, or other subcortical nuclei. In the model to be described, the fine-structure of the network has to be taken into account; Lorente de No's beautiful diagrams of the cortex (11) are invaluable for this purpose. They reveal a complex vertical organization amongst the different layers, including one conspicuous circuit to which Lorente himself draws attention. This is the return of collaterals from the long descending axons of some cells back to the region of the cell body, where they appear to end on Golgi type II, or similar short-axon cells. The Golgi cell axons proliferate in baskets around the bodies of the long-axon cells "... so that the discharge of the cell with a short axon results in powerful, practically simultaneous stimulation of a large number of other cells" (11, p. 303). It might appear from this description that the feedback circuit is admirably suited to produce uncontrolled

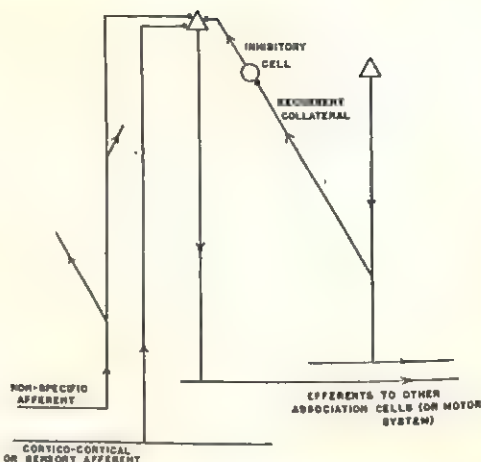


FIG. 1. Diagram of the postulated connections of cortical neurons.

epileptic discharge in the cortex, but recent microelectrode work by Eccles, Fatt, and Landgren (4) indicates that in the spinal cord, where a similar organization of long and short-axon cells is to be found, the short-axon cells inhibit the larger ones. It is likely that this is also true in the cortex.

The following highly simplified organization will therefore be assumed for the cortex (Fig. 1). Only two types of neuron are considered, those with long axons (having excitatory connections with other parts of the cortex or with the motor system), and those with short axons (which have local inhibitory connections). Those of the short-axon type receive impulses from the long-axon cells in their vicinity via recurrent collaterals, which leave the main shaft of the axon near the inner layer of the cortex and return toward the surface. Apart from those in the sensory cortex, which receive additional afferents from the sensory pathways, the long-axon cells receive their input from two major sources: (a) from other cortical cells, either directly via cortico-cortical association fibers, or indirectly by way of distributing cells in the association nuclei of the thalamus; and (b) from the nonspecific

projection system, which will here be regarded as a purely excitatory arousal system. This second source of afferent stimulation must play an important role in motivation, and it will be discussed in that connection later; for the moment, it need only be noted that we postulate that the facilitatory bombardment it provides is necessary for the conduction of impulses within the cortico-cortical network.

It appears from Lorente's diagrams (11, pp. 296, 298) that the long axons from the association cells branch and travel through the white matter to a number of different parts of the cortex. Each of these branches divides again as it ascends through the layers of the cortex to synapse with other long-axon cells. With this in mind, we can now proceed to analyze the activity in this network in the way that we previously analyzed that of Hebb's model. We shall start by assuming that no cortical cells are firing, but that there is a sufficient facilitation from the nonspecific projection system to make transmission between cortical cells possible. Suppose now that one cortical cell is somehow fired, and that it has effective connections to several other long-axon cortical cells (ten, let us say). These ten long-axon cells will in turn fire about a hundred others, and so on. At this rate the activity would soon snowball to an astronomical value, so that it clearly cannot continue to multiply for very long; in fact, the spread will quickly be checked by the firing of inhibitory cells. As the density of firing cells increases, impulses from many sources will arrive almost simultaneously in any small region of the cortex. When a long-axon cell is fired by one of the impulses, it will fire the short-axon cells in its vicinity and so prevent many subsequently arriving impulses from firing other neighboring cells. Thus, many of the impulses circulating in the cortex will be

lost because they arrive in the regions of strong inhibition surrounding recently fired cells. Finally, a level of activity will be reached at which only ten per cent of the impulses are able to fire cells, so that in the example given (in which one cell could fire ten others in the absence of inhibitory influences) a state of equilibrium will be reached. If more long-axon cells start to fire, the regions of inhibition will be crowded together and less than one tenth of the circulating impulses will fire new cells; the activity must therefore drop back to the equilibrium value. If, on the other hand, too few cells are fired, the multiplication factor will increase above unity and the total activity will increase again. It is true that many factors have been neglected in this analysis; some of them are important and will be considered in a moment; other difficulties (such as the fact that when the level of neural activity rises, the increase in inhibition will be offset to an indeterminate extent by increased temporal and spatial summation) cannot be dealt with until better quantitative data are at hand. In the meantime we can only postulate that the parameters involved have values within the limits necessary to allow the network to behave in the manner described.

For reasons that will be clear later, it is desirable to construct the model in such a way that the equilibrium activity continues to circulate within a single group of cortical cells, instead of progressing continuously to new cells at each step in the firing. To this end, the following postulate will be made regarding the anatomical organization: the short-axon inhibitory cells which receive recurrent collaterals from a long-axon cell have fewer inhibitory connections to that particular long-axon cell than they do to other long-axon cells in the region. This is illustrated in an exaggerated way in Fig. 2, where it is seen

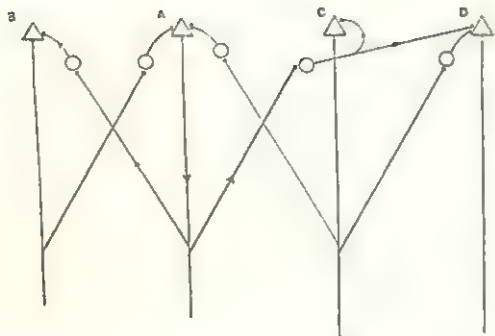


FIG. 2. Types of recurrent inhibitory connections.

that when Cell A is firing it causes the inhibition of its neighbors, B, C, and D, but is not itself inhibited. In fact, because the surrounding long-axon cells cannot now be fired, there is no way in which the short-axon cells discharging on to A can be fired. Therefore, as long as A continues to fire, it protects itself from being inhibited. This is not the only mechanism—perhaps not even the most likely mechanism—by which the desired result could be achieved (repetitive firing during the hypersensitive phase of the recovery cycle would also tend to confine the activity to the group of neurons that were already firing), but it is a useful hypothesis on which to base subsequent discussion.

Let us now see how this postulate will affect the equilibrium activity that can be sustained in our model of the cortex. There is a good chance that the cell which originated the activity will receive an afferent impulse from some other cell which is fired later in the avalanche. If it does, there will be nothing to prevent its firing again in response to the returning impulse, because its original discharge will have inhibited neighboring cells, thus greatly reducing the probability that it will be inhibited itself. When it fires again it will also be able to re-excite the ten cells that it fired previously, because they will have likewise remained free

from inhibition. Thus the whole group will tend to re-excite itself, confining the activity to those neurons in which it first spread. If by some chance the original cell does not receive connections from any of the cells subsequently fired, some of the next ten cells, or those further along in the chain of activity, will do so, and the final result will be much the same.

It now remains to see what prevents the activity from persisting in this group of cells indefinitely. A property of most cells in the nervous system is that if they are subjected to a constant source of excitation their frequency of discharge steadily falls. We may therefore expect a decline in the intensity of firing in the group. As each cell adapts, it will deliver fewer impulses to other cells in the group, and reduced bombardment will combine with the adaptation to slow down the activity. On the other hand, there will be a concomitant decrease of inhibitory firing, so that the over-all activity will not decline as rapidly as it otherwise would. However, there is a limit to the compensation that can be provided in this way, because when the inhibition becomes too small some of the cells outside the group, which were previously prevented from firing, will be released. The firing of these cells will then bring in fresh short-axon cells which will immediately inhibit any fatigued cells in the vicinity. Thereupon, the balance of excitation and inhibition will be disturbed in regions to which the affected long-axon cells send impulses. If most of the cells in the group were already becoming adapted, or fatigued, the disturbance will snowball in reverse; each cell that stops firing causing connected cells to stop also. In this way the firing of one group might come to an abrupt end, and another group of cells take over. Note that although the activity will swing from one group of cells to another, it



can never die out; long before that happened the level of inhibition would become so low that all the unfatigued cells would be able to fire. The self-maintained cortical activity described above is, as the reader has no doubt realized, intended to form the basis of the revised version of the cell assembly.

One behavioral phenomenon which Hebb explains in terms of the cell assembly is "set." In order to do the same with the present model we shall have to add a further neural postulate. Perhaps the simplest type of behavior exemplifying "set" is that in which a stimulus, A, followed by another stimulus, B, gives rise to the response, C, although neither A nor B alone will produce the response. Presentation of A induces a set to respond to B. In neural terms, it seems that after a cell assembly has ceased to exert a direct control on an animal's behavior, it continues to influence the way in which that behavior develops, perhaps for a period of minutes. Only two categories of neural process seem able to account for this long-lasting aftereffect of a cell assembly; either some neurons of the group representing Stimulus A, for example, continue to fire, though the main body of the activity is now taking place in cells of the group representing Stimulus B (this is the "active" trace hypothesis), or the activity A leaves behind it physical or chemical changes which last long enough to influence subsequent firing patterns (the "latent" trace hypothesis). Hebb attempts to explain "set" in terms of the first of these alternatives—the active trace hypothesis—and the same process might be applicable to the present model. However, the second mechanism seems to offer much more interesting possibilities, which we shall now proceed to explore. The hypothesis might be formally stated as follows: If afferent impulses fall on a cortical cell but do not fire it (because of inhibition

or lack of adequate background facilitation), they nevertheless leave the cell with a lowered threshold, the effect dying away with a time constant of many seconds. This influence of one cortical neuron on another will be called "priming." This postulate complicates the original model slightly because now, at the moments of transition from one pattern of cortical activity to another, we have to take account of the residual priming left by earlier activities, as well as of the momentary excitations. For example, a neuron that receives synaptic connections from neurons of several previously active groups will have a better chance of firing than a neighboring neuron that has been stimulated only by neurons of the most recently active group.

We must now examine the effect of sensory afferents on this postulated cortical system. The sensory projection areas of the cortex are regarded for this purpose as the final distribution centers for sensory impulses. The long-axon cells with specific-sensory afferent connections are assumed to have widespread efferent connections to the rest of the cortex, so that a change in the pattern of sensory input will result in multiple changes of the local fields of excitation there. If the cortex is already intensely active, and if the sensory change is a slight one, the added excitation will be unable to affect the ongoing pattern of cortical firing. The cells which are not already firing will be too strongly inhibited by those that are firing to be influenced by the input. Behaviorally, this means that the animal will ignore the stimulus and pursue its former course. On the other hand, if the stimulus is a strong one, or if the cortex is relatively inactive when it is presented, the sensory impulses will determine, or significantly influence, the pattern of the next activity in the cortex. The learning mechanism can now be

made explicit. It has been postulated that if a neuron helps to fire another neuron on several occasions, it becomes a more potent agent for firing that neuron subsequently. This clearly implies that the neurons of a group that fire one another, in the way that has been described, will tend to bind themselves together more closely. Let us assume that one group has been activated on several occasions, and that consequently each neuron in the group has strengthened its synaptic connections with one or more other cells in the group, but not with any of the nonfiring cells to which its connections were initially equally strong. At this stage, if a neuron in the group is fired, the likelihood of its firing other cells of the group will have been increased; and the more cells of the group that are so aroused, the more the remaining members will be subjected to preferential excitation. This selective effect will often cause the total group activity to be established, starting from the firing of only a fraction of the constituent neurons. Moreover, each time this concerted firing takes place, the connections will be further strengthened, and the probability of the group's firing as a unit will become still greater. This process is equivalent to Hebb's mechanism for building up a cell assembly by repeated presentation of a "perceptual element"; its significance for the explanation of perceptual learning and concept formation has been discussed in some detail by Hebb (5, 6) and need not be repeated here. It must not be supposed, however, that after a stimulus has been presented often enough it will become able to determine the total pattern of the assembly; there will always be an inconstant "fringe" of cells that fire because of some previous cortical activity, metabolic conditions, or other factors independent of the stimulus. This fringe plays an important role in the associa-

tion of one cell assembly with another, as we shall now see.

Associative learning was mentioned earlier as being the linking together of cell assemblies into phase sequences. This will usually occur because the assemblies in question have been aroused contiguously. It has been assumed up to now that the neurons that initiate an assembly activity are themselves fired by a sensory event, but clearly once a group of cells has strengthened its interconnections sufficiently, it does not matter which of its elements are fired first; the activity will in any case spread to the rest. If enough cells of one assembly are left in a primed or excited state by firing in other cell assemblies, that assembly may burst into activity when a previous one fades out. The question we have to answer is how such an interrelationship between two assemblies could be established by experience.

Let us suppose that two stimuli, giving rise to Cell Assemblies A and B, respectively, are presented in succession a number of times. Under these conditions some of the fringe neurons of Assembly B will be determined by A, and so will not be random any more. In other words, some of the cells primed by A will be active with Assembly B on every occasion that the stimulus pair is presented, and these cells will start to acquire stronger connections with the cells of the B assembly. Before long, in fact, they will become members of the assembly. When this has happened it will be possible for these cells, which on the first presentation of the stimulus pair were merely a part of the random fringe of B, to arouse the whole B cell assembly in the same way that any other fractional firing of the assembly would arouse it. It remains only to point out that these particular cells are the ones that are so strongly primed by the activity of the A assembly that they have a high probability of firing when



it stops (which is how they became part of the firing of B in the first place), and the method by which two assemblies become associated will be evident. Of the cells released from inhibition when Assembly A dies away, those that have on previous occasions fired with (and thus acquired strong connections with) cells of the B assembly, will be able to recruit the rest of the cells of that assembly. On the other hand, the successors of Assembly A which are incompatible with firing in the B assembly will have no strengthened connections, and will be promptly suppressed as the rest of the B assembly begins to fire.

It may be observed that the amount of perceptual "overlearning" that has gone into the development of the assembly, before its association with another, will make a great deal of difference to the ease with which the linking is achieved. If the total assembly action can spring from the firing of, say, five per cent of its constituent cells, links will have to be strengthened to a smaller number of cells than would be necessary if twenty per cent of the cells had to be fired to insure assembly action. That is, if the intragroup connections are weak to start with, the new cells will have to acquire stronger and more numerous connections to the rest of the group in order to compensate for the deficiency. Another point to be remembered is that only rarely will a single cell assembly be able to fire another one unaided, as in the example given. In most cases, the assembly aroused by association will have been primed by many previously active assemblies before it builds up enough sensitized cells to enable it to fire without its proper sensory input. This is an explanation of the importance of context, and set, for association; the arousal of a particular association may depend on the surroundings being similar to those in which the original learning took place, because the

firing of a number of cell assemblies by the environmental stimuli is necessary to build up priming in the required group.

The model described above represents a brain mechanism capable of associating contiguously presented stimuli. This is only part of the apparatus needed to explain an animal's behavior in a learning situation, but it is an essential part. The most important other mechanism has to do with motivation. In this paper motivation will be considered to influence learned behavior in two ways: (a) by affecting the rate of learning, and (b) by helping to evoke responses.

In order to discuss motivation from a physiological standpoint, we must return to the system which was mentioned briefly before, the nonspecific projection system. It will be remembered that our model of the functional organization of the cortex was developed on the assumption of a fixed level of facilitatory bombardment from this system, such that one long-axon cortical cell was able to fire ten others in the absence of inhibition. If in the illustration we had assumed a lower level of nonspecific activity, the total cortical firing would have been less, though not proportionally so because of the accompanying decrease in inhibitory firing. It might be remarked, in passing, that a number of loose ends may be rather neatly tied together if it is assumed that impulses reaching a long-axon cell in the cortex from other cortical cells do not actually fire that cell, but induce a long-lasting depolarization that sensitizes it to bombarding impulses from the arousal system. Thus the cell would be *primed*, but not fired, by other cortical firing, and then fired by the next few impulses that it receives from the nonspecific arousal system. Such a mechanism would be consistent with the observation of Li and Jasper (9) that the rate of cortical firing increases with the level



of arousal; it would also be in agreement with the data of Clare and Bishop (2) on the long-lasting dendritic depolarization set up by association afferents. Moreover, the work of Bradley (1) might be interpreted as indicating that the arousal system afferents are cholinergic, and as acetylcholine is destroyed instantly on being released it is unlikely that the arousal system impulses would give rise to any long-lasting effect in the cortex.

Lindsley (10) has indicated that there is a rough correlation between the motivational and emotional states of an animal and the level of activity of its arousal system; and Sharpless (12) has shown that novel stimuli are more potent stimulators of the arousal system than are familiar ones. It is also generally true that a hungry or otherwise uncomfortable animal has an increased level of arousal activity, especially in the presence of a satisfier. It thus seems that conditions which increase firing in the nonspecific system are those which tend to promote learning. There is still dispute as to how essential reinforcement might be for learning, but there is little doubt that some sort of "emphasis" on the relevant stimuli accelerates the learning process. In terms of our model, this means that cortical cells are linked together more rapidly when the rate of cortical activity is higher—a most reasonable conclusion. It might also be deduced from the model that *some* association should be possible as long as *any* cortical cells are firing (during dreams, for example). The reinforcement controversy thus reduces to the question of *how much* firing is going on in the cell assemblies concerned, reinforcing agents serving to increase that firing through their effects on the nonspecific projection system. According to this theory, therefore, the arousing rather than the need-reducing properties of the reinforcement are those important for

learning. Evolution would account for the fact that, for most present-day animals, need reducers and dangerous situations tend to be particularly arousing.

It should be noted that, in the examples given, it has not been a response but rather a relationship between stimuli that has been learned. It is not too difficult to extend the model to account also for the association between stimuli and responses, but we then have to face the problem of why the stimulus is sometimes followed by the response and at other times not. No doubt this is sometimes due to the failure of effective association between the cell assemblies concerned, but we have introspective evidence which suggests that we may know what to do to achieve a goal, but not be sufficiently interested to perform the actions. This brings us to the problem of the effect of motivation on performance. Unfortunately, we cannot do justice to this topic in the space of this paper, but a brief indication will be given as to how the model might be able to handle it. We must assume that many motor neurons will be fired along with the neurons of the cell assemblies we have described, but because they must compete with postural and other reflexes for control of the musculature, they can only produce a response if they fire at a high enough frequency. Below some threshold value, cortical firing has no effect on overt behavior. It is suggested that the arousal system is responsible for raising cortical activity above this threshold when the stimulus conditions are sufficiently exciting. (However, as Hebb has suggested [7], if the firing becomes too vigorous, some other factor may intervene to prevent further responses.) In the case of motivation set up by expectancy of reward or punishment, it must be assumed that the arousal system has been excited, not by the stimulus actually present, but by a cell assembly representing the moti-

vating stimulus which, through previous experience in the situation, had become associated with the cell assembly for the actual stimulus. The cell assembly for the motivating stimulus will be able to fire the arousal system because of an association built up through innumerable presentations of the motivating stimulus, under conditions where it fired both the cortical and the subcortical systems simultaneously. Thus the environmental stimuli fire their cell assemblies, these arouse the assembly of the motivating stimulus by association, and this in turn fires the arousal system, also by association.

An example might make the proposition clearer. When a rat is first placed in a maze, the arousal value of the novel surroundings will be sufficient to raise the level of cortical firing so that more or less randomly associated "motor" cell assemblies will influence overt behavior; the rat will explore. If no incentive is introduced, the effect of novelty will wear off and the rat will lie down and doze. However, if the rat is food-deprived, and food is placed in the maze, the interaction of the stimuli from the food with the cells in the rat's nervous system that are sensitized by deprivation will produce persistent excitation of the arousal system. Moreover, cell assemblies fired by the food stimuli (if the food is familiar) will have strong connections with the cells in the arousal system, so that even if the deprivation is not severe the presence of food will be likely to produce arousal. These cell assemblies for food stimuli will, of course, become associated with any other cell assemblies active at about the same time (those representing the sight of the food dish, or the feel of the maze floor, for example). Therefore, when on a subsequent occasion the hungry rat is returned to the maze, one of the assemblies fired by the stimuli in the starting box may be able to excite, by associa-

tion, the assembly which represents the food, and thus, indirectly, the arousal system. The firing in the latter will not die out as soon as the cell assembly which aroused it ceases to fire; activity in the arousal system does not stop abruptly (possibly because its activity releases adrenalin into the blood stream). It will therefore continue to bombard the cortex at an increased rate for perhaps several minutes, and the motor component of any cell assemblies occurring during that time will fire intensely enough to produce a response. If, for example, a running response has been associated with the starting-box stimuli, the rat will actually run; if the same cell assembly (for running) had become active before the arousal system had been fired by the association between the starting-box stimuli and food, then the cortical firing would probably not have been intense enough to elicit overt movements. The same sluggishness would be expected if the animal was sated, so that the cells of the arousal system were not sensitive to food stimuli or their associated cell assemblies.

The above example is admittedly superficial. More complicated situations, such as avoidance learning, or extinction, cannot be explained without introducing still more postulates, in particular a system for inhibiting movements. The value of such a model would be primarily to demonstrate that it is possible to design a machine that would not only "want" or "avoid" arbitrarily designated stimuli, but also learn what to do about them under all environmental conditions to which it was sensitive. Whether the design would bear any relationship to the actual mechanism of animals which can also do these things could be known only after further research.

#### SUMMARY

A neural model has been presented, based on an "association-of-ideas" para-

digm of learning, similar to that used by Hebb. In it groups of neurons (cell assemblies) become representors of stimuli, and can then be linked together by being fired contiguously. The model differs from Hebb's in that an inhibitory regulatory system is postulated which limits (to a minute fraction of the total) the number of cortical neurons that can fire simultaneously, and insures that those firing are dispersed as widely as possible. A further change is introduced to meet the paradox that cell assemblies can be associated with one another without losing their individuality and being submerged in a composite new cell assembly. In association, it is not the cells of one assembly that acquire connections with the cells of another; instead, cells primed, or sensitized, by the first assembly become incorporated into the second. Thus, one principle of learning—the binding of cells into a group by repeated simultaneous firing—fulfills a double role; when the newly added cells are predominantly primed by sensory input perceptual learning results; and when the new cells are primed by the firing of another cell assembly, associative learning results.

Because the effect of priming lasts for many seconds, it is possible for a cell assembly to accumulate the sensitizations induced by the activities of a number of associated assemblies, and so increase the probability that it will itself fire.

The dual role of motivation—the facilitation of learning and the elicitation of responses—has been discussed in terms of the arousal effects of the non-

specific projection system on the postulated cortical network.

#### REFERENCES

1. BRADLEY, P. B. The effect of some drugs on the electrical activity of the brain in the cat. *EEG. clin. Neurophysiol.*, 1953, 5, 471. (Abstract)
2. CLARE, M. H., & BISHTOP, G. H. Dendritic circuits: the properties of cortical paths involving dendrites. *Amer. J. Psychiat.*, 1955, 111, 818-825.
3. ECCLES, J. C. *The neurophysiological basis of mind*. Oxford: Clarendon, 1953.
4. ECCLES, J. C., FATT, P., & LANDGREN, S. Central pathway for direct inhibitory action of impulses in largest afferent nerve fibres to muscle. *J. Neurophysiol.*, 1956, 19, 75-98.
5. HEBB, D. O. *The organization of behavior*. New York: Wiley, 1949.
6. HEBB, D. O. The problem of consciousness and introspection. In *Brain mechanisms and consciousness*. Springfield: Thomas, 1954.
7. HEBB, D. O. Drives and the C.N.S. (Conceptual nervous system). *Psychol. Rev.*, 1955, 62, 243-254.
8. HEBB, D. O. A neuropsychological theory. In S. Koch (Ed.), *Psychology: a study of a science*. Vol. 1. McGraw-Hill, in press.
9. LI, C. L., & JASPER, H. Microelectrode studies of the electrical activity of the cerebral cortex of the cat. *J. Physiol.*, 1953, 121, 117-140.
10. LINDSLEY, D. B. Psychological phenomena and the electroencephalogram. *EEG. clin. Neurophysiol.*, 1952, 4, 443-456.
11. LORENTE DE NÓ, R. Cerebral cortex: architecture. In J. F. Fulton, *Physiology of the nervous system*. (3rd ed.) New York: Oxford Univer. Press, 1949.
12. SHARPLESS, S. K., & JASPER, H. Habituation of the arousal reaction. *Brain*, 1956, 79, 655-680.

(Received November 28, 1956)



## LEVEL OF ASPIRATION AND DECISION MAKING<sup>1</sup>

SIDNEY SIEGEL

*Center for Advanced Study in the Behavioral Sciences<sup>2</sup>*

The purpose of this paper is to discuss two related topics: (a) the role of a person's level of aspiration in his decision making, and (b) the measurement of level of aspiration in a decision-theory context. The discussion of the second topic will include a brief summary of some experimental evidence which lends support to the methodological stand taken here.

The notion of level of aspiration is invoked in reference to the goal-striving behavior of an individual when he is presented with a task whose outcome can be measured on an achievement scale. Level of aspiration refers to the particular achievement goal for which the person strives.

The concept of level of aspiration was introduced by Dembo (8), and the first experiment in the area was conducted by Hoppe (14). It is a familiar concept to psychologists and others, having been the topic of extensive discussion and experimentation in the last quarter of a century. An early review of the literature is given by Frank (13). Rotter has offered a critical review of the methodological aspects of level of aspiration studies (19). An exceptionally important theoretical article, by Lewin, Dembo, Festinger, and Sears, appeared in 1944 (16).

<sup>1</sup> This paper is based on research conducted under Research Grant M-1328 from the National Institute of Mental Health and under a grant from the Council on Research of The Pennsylvania State University. The author is indebted to Selwyn W. Becker, Danuta M. Hiz, and Henry T. Hiz for discussing with him many of the ideas presented here.

<sup>2</sup> On leave from the Pennsylvania State University, 1957-1958 year.

### DECISION THEORY AND LEVEL OF ASPIRATION THEORY

It is the contention of the present writer that the psychological situation established in level of aspiration experiments may profitably be characterized as a *decision* situation, for from the alternative possible goals the individual must decide for which goal he will strive.

It is a remarkable fact that, by a simple change in nomenclature, the theoretical model used by Lewin *et al.* (16) in the prediction of the choices (decisions) of individuals in a goal-striving situation—a model based on the work of Escalona (11) and Festinger (12)—may be rendered fundamentally equivalent to the theoretical model employed by decision and game theorists. This latter theory, first advanced by Bernoulli (2), discussed by Ramsey (18), and formalized by Von Neumann and Morgenstern (28) and by Savage (20), states that under conditions of uncertainty individuals behave as if they were attempting to maximize expected utility. According to these and other decision theorists, an individual's decisions underlying his choices among alternatives involving uncertain outcomes (outcomes with stated probabilities of attainment) are based on the utilities of the entities (objects, actions, goals, etc.) and on the probabilities (subjective probabilities, for most decision theorists) associated with attainment of the entities. The decisions are a function of these two variables (utility and subjective probability) in that the individual seeks by his choices to maximize the sum of the products of probability and utility, i.e.,

he acts so as to maximize

$$SEU = \sum_i p_i u_i,$$

where *SEU* is subjectively expected utility (cf. [10]).

*Equivalence of terms.* Lewin *et al.* assert that the level of aspiration set by the person is a function of three factors: (a) "the seeking of success," (b) "the avoiding of failure," and (c) the "cognitive factor of a probability judgment" (16, p. 376). The last-mentioned factor refers to the individual's subjective probability of achieving a goal. (Lewin *et al.* use the term "subjective probability." See [16] p. 361 *et seq.*) Each of these three concepts of Lewinian theory may be "translated" directly to an equivalent concept in decision theory. The achievement scale may be thought of as a utility scale, where *utility* refers to the subjective value of an entity (goal, action, object) as distinguished from its objective or "face" value. It is the contention of this paper that each level on an achievement scale has a certain utility for a person, and that therefore level of aspiration may be conceived as a "point" on a scale of utility. An achievement which represents "success" in Lewinian terms may be thought of in decision theory terms as one which has "positive utility." An achievement which represents "failure" to the person may be thought of as having "negative utility" for him. And the Lewinian notion of "subjective probability" needs no translation, for the same term is used in decision theory.

The demonstration that certain concepts in Lewinian level of aspiration theory are equivalent to certain key concepts in decision theory is important because of the consequences it may have. Because of these equivalences, two important bodies of research literature may be synthesized, with benefit to both.

The benefit to traditional psychologists, accustomed to dealing with terms like "success" and "failure," will accrue because of the much greater formality and rigor characteristic of decision theory, a rigor unknown to many psychologists. The benefit to decision theory will come because the synthesis will render available to decision theorists a large amount of existing experimental evidence with relevance to their work which has not yet been recognized by them.

*Equivalence of formal statements of relations.* The similarity between the two theories does not rest with equivalence of their variables. Also comparable are the formal statements of the relations among these variables.

For example, Lewin *et al.* (16, p. 360) state that:

The individual faces the possibility of succeeding or failing and the positive or negative valence of such a future success or failure on the various levels is one of the basic elements for the decision. To determine the valence ( $Va$ ) of each level ( $n$ ) of activity [ $Va(A^n)$ ] we have to consider the negative valence that future failure has on that level [ $Va(Fai A^n)$ ] and the positive valence of success on that level [ $Va(Suc A^n)$ ].

$$Va(A^n) = Va(Suc A^n) + Va(Fai A^n).$$

When discussing the individual's felt probability of success (16, p. 362), they also state that:

The most simple assumption is that this "weighted" valence of success [ ${}^oVa(Suc A^n)$ ] is the product of the valence and the probability of success

$${}^oVa(Suc A^n) = Va(Suc A^n) \cdot Prob(Suc A^n).$$

The corresponding formula for failure is:

$${}^oVa(Fai A^n) = Va(Fai A^n) \cdot Prob(Fai A^n).$$

If Lewin *et al.* had proceeded to the logical conclusion of these ideas, they could have stated that

$$Expected Va(A^n) = Prob(Suc A^n)$$

$$\cdot Va(Suc A^n) + [(1 - Prob) \\ \times (Suc A^n)][Va(Fai A^n)],$$

which in decision theory terms is given as

$$\text{Expected Utility} = \text{Probability} \cdot \text{Utility of Success} + (1 - \text{Probability}) \cdot \text{Utility of Failure}$$

in which it is asserted that the individual chooses among alternatives (sometimes called "courses of action") so as to maximize expected utility.

This formula is equivalent to that given by Von Neumann and Morgenstern (28), and has been utilized in studies of decision behavior by Davidson, Siegel, and Suppes (7), Mosteller and Nogee (17), Siegel (21), Hurst and Siegel (15), Edwards (10), and others (see [9]).

In the discussion of Lewin *et al.*, the problem of obtaining a rigorous measurement of utility (degree of valence) and of subjective probability is virtually disregarded, despite the fact that these are the variables upon which the entire structure of the model depends. In the literature of decision theory, on the other hand, attention has been given to this important facet. See, for example, the experiments by Mosteller and Nogee (17) and by Davidson, Siegel, and Suppes (7) on the measurement of utility in an interval scale and on the measurement of subjective probability, and also those of Siegel (21) and Coombs and Komorita (6) on the measurement of utility on an ordered metric scale.

Another contrast between the two theories rests on the fact that in decision theory the formal theoretical models are based on a set of assertions from which consequences can be derived by logico-mathematical arguments. The assertions are referred to as axioms of the system (cf. [20]). If the model is used to interpret "real world" phenomena, it is possible to know, without ambiguity, what aspect of the model is analogous to what thing (or action) in the real world.

When such an abstract and formal system is available, it sometimes happens that the mathematical system will serve as a model for more than one situation in the real world. This fortuitous event may reveal analogies or relations of whose existence scientists were previously unaware. This seems to be the case in the present instance. A formal theoretical model has been constructed which is useful in interpreting decision behavior in general. It is contended in this paper that this model will encompass decision behavior in goal-striving or achievement situations, i.e., in situations where level of aspiration is an important variable of behavior, and furthermore that level of aspiration may usefully be defined in decision-theory terms and measured in a decision-theory context.

#### LEVEL OF ASPIRATION AS A DETERMINANT IN DECISION MAKING

Suppose an individual is given the task of finding as many as he can of five hidden objects, which will be his to keep. Suppose further that this individual aspires to find and thereby acquire at least four of these. That is, the acquisition of four or more of the objects is defined as "success" by this individual, and the acquisition of any fewer will constitute "failure" for him. Now suppose that before he attempts to find the objects, he is asked to choose between the following alternatives: (a) a 50-50 chance of getting either four of the objects or none of the objects, or (b) assurance of getting three of the objects. That is, suppose that the outcome of the task is completely determined by the individual's choosing either of the alternatives above and then by the toss of a "fair" coin. If the individual chooses Alternative (a) above, and if the coin lands "heads" he gets four objects, whereas if it lands "tails" he gets none at all. On the other hand,



if he chooses (b) he gets three objects if the coin lands at either heads or tails.

Now since the individual has set the acquisition of four objects as his level of aspiration, and the acquisition of any fewer will be considered as a "failure" by him, it seems clear that he will choose Alternative (a), despite the fact that objectively Alternative (b) has greater expected value. In other words, Alternative (b) may have greater expected *value*, but Alternative (a) has greater expected *utility* for this person. His level of aspiration has made the utility of the objects nonlinear with their objective value.

In game-theory terms, the above situation may be described as a one-person game in which the subject chooses between two alternatives, each of which is a probability-combination of two outcomes. The format for the game is:

	Alternative (a)	Alternative (b)
If heads occurs	you get 4	you get 3.
If tails occurs	you get 0	you get 3.

The subject chooses the column; the outcome of the toss of the coin determines the row.

Now if the subject has Level of Aspiration 4, and if therefore he chooses Alternative (a) over (b), this tells us that for this individual the difference in utility between 4 and 3 objects is greater than the difference between 3 and 0.

That conclusion is based on this reasoning: if

$$\begin{aligned} &\text{Expected Utility of (a)} \\ &> \text{Expected Utility of (b),} \end{aligned} \quad (1)$$

where  $u(x)$  is read as "the utility of acquiring  $x$  objects" and interpreted to mean the subjective value of  $x$ , i.e., its worth to the person, then Statement 1 can be written

$$\begin{aligned} p \cdot u(4) + (1 - p) \cdot u(0) &> \\ p \cdot u(3) + (1 - p) \cdot u(3). \end{aligned} \quad (2)$$

If  $p$  is understood to be subjective probability, and is known to be one-half, then Statement 2 may be written

$$u(4) + u(0) > u(3) + u(3) \quad (3)$$

and

$$u(4) - u(3) > u(3) - u(0). \quad (4)$$

That is, when Alternative (a) is preferred, we may infer that Statement 4 above is true, i.e., that the difference in utility between 4 and 3 objects is larger than the difference in utility between 3 and 0 (21, pp. 212-213).

We can see from the above that for this individual the acquisition of four or more objects has positive utility. The acquisition of three or fewer has negative utility. Zero utility is represented by some point between three and four. It should also be noticed that *the level of aspiration is associated with the largest difference on the utility scale.*

We may cite another example to illustrate the same points. Suppose a student is enrolled in a course in which he aspires for the grade of B. Before the examination on which the course grade is based, he is asked to choose between the following alternatives:

	Alternative (a)	Alternative (b)
If Event E occurs	you get an A	you get a B.
If Event Non-E occurs	you get a C	you get a B.

If the individual believes that the probability of the occurrence of Event E is .50, it seems clear that he will choose Alternative (b), since his level of aspiration is to achieve a grade of B. By this choice, the individual demonstrates to us that the difference in utility between B and C is greater than the difference in utility between A and B. That is, on his utility scale B is closer to A than to C. See Fig. 1.

	A	B	C
FIG. 1.			
Now the same person is asked to choose between these:			
	Alternative (a)		Alternative (b)
If Event E occurs	you get a B		you get a C.
If Event Non-E occurs	you get a D		you get a C.

If his subjective probability toward Event E is .50, he will choose Alternative (a), for in terms of his level of aspiration any grade less than B is a "failure." His selection of Alternative (a) demonstrates that for him the difference in utility between Grades B and C is larger than that between C and D, as shown in Fig. 2. Again we see that the level of aspiration is associated with the upper bound of the largest distance on the utility scale.

### LEVEL OF ASPIRATION DEFINED

Before we offer a formal definition of the level of aspiration, we may summarize what has been said above concerning it:

1. The level of aspiration of an individual is a position on his utility scale of an achievement variable.

2. The lower bound of that position is neutral in utility, i.e., all points below the lower bound have negative utility (psychological feeling of dissatisfaction) associated with them and all points above the lower bound have positive utility (psychological feeling of satisfaction) associated with them.

3. The level of aspiration is associated with the upper bound of the largest distance on the utility scale.

	B	C	D
FIG. 2.			

Each of the characteristics of level of aspiration enumerated above is incorporated in the following formal and more rigorous definition of level of aspiration:

*The level of aspiration of an individual is a point in the positive region of his utility scale of an achievement variable; it is at the least upper bound of that chord (connecting two goals) which has maximum slope; i.e., the level of aspiration is associated with the higher of the two goals between which the rate of change of the utility function is a maximum.*

It may be seen from this definition that the difference in utility between achieving at the level of aspiration or achieving at the next lower level is greater than is the difference in utility between achieving at the level of aspiration or achieving at the next higher level. In other words, the level of aspiration is that goal which has the

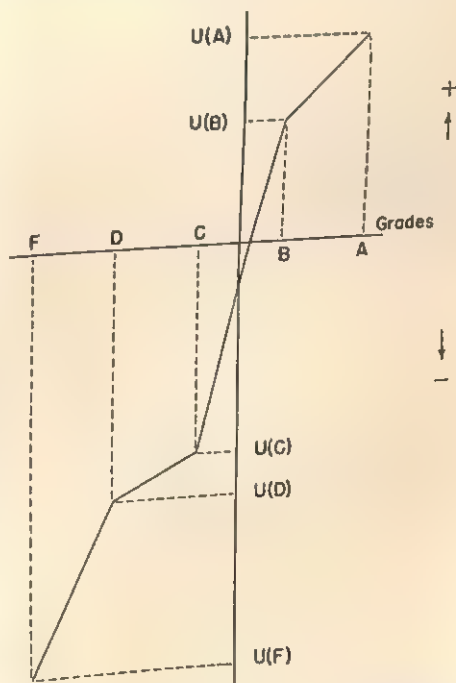


FIG. 3. Graphical representation of the utility of grades for an individual whose level of aspiration is "B."

A	B	C	D	F
---	---	---	---	---

FIG. 4.

largest distance (difference) in utility between it and the next lower goal.

Figure 3 demonstrates this graphically. The curve represents the utility function of a student who aspired for a B and who made the choices discussed above, among others. It may be seen that the differences between the utility of B,  $U(B)$ , and the utility C,  $U(C)$ , is larger than any of the other differences. The ordered metric utility function may also be represented as it is shown in Fig. 4.

#### LEVEL OF ASPIRATION AS A PROBLEM IN MEASUREMENT

The preceding discussion is directed to the point that under certain conditions the task of ascertaining an individual's level of aspiration reduces to the task of measuring his utility of the goals in question. That is, when an achievement variable is involved, the level of aspiration on the achievement scale may be derived by ascertaining the individual's utility function for the goals in question.

How strong must the measurement be? Obviously the weaker sorts of measurement—nominal and ordinal scaling—will not be satisfactory, for such scaling yields no information about the distances between the entities on the scale. Since the identification of level of aspiration requires some knowledge of the distances on the scale, utility must be measured on a stronger scale than an ordinal one.

The strength of measurement which is required depends on the nature of the variate—on whether the possible achievement goals are represented by a continuous variate or are a discrete set of entities with an underlying continuum.

If the possible goals are on a continuum and if the achievement scale is represented by a continuous variate, i.e., if the scores (goals) may have any value in a certain range, then at least *interval* measurement of utility is necessary to derive level of aspiration. In interval measurement, the distances between points (i.e., the intervals) are known numerically (22, 26, 27), and thus the numerically largest distance may be identified and the level of aspiration thereby located.

If the possible goals are discontinuous and if the achievement scale consists of a discrete set of entities, i.e., if the scores (goals) can take only isolated values, then at least *ordered metric* measurement is necessary to derive level of aspiration. In strength, an ordered metric scale falls between an ordinal scale and an interval scale. An ordered metric scale (4, 5, 21) not only gives a ranking of the entities involved but also a ranking of the distances between the entities. With an ordered metric scale we can identify the largest distance and, inasmuch as the individual's choice is limited to the discrete set of goals available, we thereby find that goal which best represents his level of aspiration of those available.

The problem of obtaining an interval measure of utility has been dealt with by Davidson, Siegel, and Suppes (7). Their paper includes an axiomatization of the interval measurement of utility as well as a report of an experiment in which interval scales of utility were successfully derived for 15 subjects.

A method for obtaining an ordered metric measure of utility has been developed by Siegel (21). Examples of applications of this method in tests of various hypotheses are given in (15), (23), and (24).

Most level-of-aspiration situations, if not all, involve a discrete set of goals. Even if there is continuity underlying



the achievement scale, as is the case in achievement in throwing darts at a board with a bull's-eye, the possible goals are usually broken up into a discrete set. In the case of the dart board, these are concentric circles around the bull's-eye. Thus it would seem that in most situations measurement in an ordered metric scale is all that is necessary to derive an individual's level of aspiration.

To test the validity of the statement that the level of aspiration is that goal which is the least upper bound of the largest distance on an ordered metric scale of utility, a study was conducted which is reported fully elsewhere (1). A summary report of it is included here to lend greater clarity to some of the topics discussed.

#### EXPERIMENTAL VERIFICATION

Becker and Siegel (1) have conducted a study wherein they tested the hypothesis that the ordered metric utility function would yield information on the individual's level of aspiration.

Subjects were 20 students enrolled in an elementary course in statistics. These were students who volunteered to gamble with the instructor for their midterm grade, with the understanding that the grade they obtained in the gambling session would be entered in the course records in lieu of their score on the regularly scheduled midterm examination.<sup>3</sup> An ordered metric scale of the grades A, B, C, D, and F was derived for each subject by the method described elsewhere (1, 21). This method derives the subject's scale from his choices among alternatives like those exemplified above.

<sup>3</sup> After all data had been collected, the nature of the experiment was explained to all participants in individual interviews with them, and all agreed to take the examination and receive a grade based on their performance on it.

The subjects were required to make choices between alternatives in a number of offers. Each of the offers was numbered, and the subjects understood in advance that the offer on which the pay-off (midterm grade) would be based would be selected at random from all of the offers given. Therefore each decision they made had an equal likelihood of being the crucial one on which their midterm grade depended (cf. 21, p. 214; 20, p. 29).

After the subjects had made all the required choices, two ruses were introduced. First, the allegedly random device by which each student's grade would be determined was actually controlled so that each subject "won" a C for his grade. Some subjects responded to this pay-off with considerable dismay and disappointment. In response to this, the experimenters perpetrated the second ruse. The group of subjects was told that if they were dissatisfied with the grade they had won, they could obtain an interview with the instructor and the other experimenter, in which perhaps some way of raising their grade could be worked out. The implication contained in this announcement was that performance of extra work would be the mode of raising the grade. The subjects were told that any who desired such an interview would have to wait until the instructor placed an urgent long-distance telephone call, after which he and the other experimenter would return to conduct individual interviews with those students who desired them. The instructor said he would return in about five minutes, but actually his absence was unexplainedly extended to an hour's time before he returned for individual interviews. As will become clear, this ruse permitted a test of one of the hypotheses of the study.

The interviews were structured to obtain a careful independent measure of each subject's level of aspiration for

his midterm grade. Both experimenters were present, and their independent judgments of each subject's level of aspiration were in very close agreement ( $r_s = .99$ ). The interviews were conducted before either experimenter had any information concerning any subject's ordered metric scale of utility for grades.

The experiment was designed to test two hypotheses. For this summary it may suffice to state these briefly and to report the outcome of each hypothesis test.

Hypothesis I was that those subjects who would not wait for an interview with the instructor would be persons on whose ordered metric scales the largest distance was between D and F—persons whose ordered metric scales would reveal that their level of aspiration was below the C which they had won. This hypothesis was supported by the data. Four subjects left the room before the experimenters returned from their hour-long errand. All had ordered metric scales with the largest distance between D and F. Seemingly they were satisfied with the C grade they had been assured and therefore saw no reason to seek an interview.

Hypothesis II was that subjects' levels of aspiration, as determined from their ordered metric scales of utility, would be positively correlated with their levels of aspiration as judged from the interview material. This hypothesis was confirmed: the correlation between the two independent indices of level of aspiration was  $r_s = .83$ .

#### GOALS FOR FUTURE RESEARCH

Child and Whiting (3, p. 508) have formulated five general statements which contain the conclusions which may reasonably be drawn from research on level of aspiration to date:

1. Success generally leads to a raising of the level of aspiration, and failure to a lowering.

2. The stronger the success, the greater is the probability of a rise in level of aspiration; the stronger the failure, the greater is the probability of a lowering.

3. Shifts in level of aspiration are in part a function of changes in the subject's confidence in his ability to attain goals.

4. Failure is more likely than success to lead to withdrawal in the form of avoiding of setting a level of aspiration.

5. Effects of failure on level of aspiration are more varied than those of success.

Each of these generalizations may be treated as hypotheses to be tested in experimentation employing the decision-theory approach outlined here. If such experimentation leads to confirmation of the hypotheses, this would strengthen confidence in the validity of that approach.

Simon (25, p. 110) points out that in most models of rational choice it is assumed that the individual will evaluate the outcomes of all possible alternatives before he makes any choice. In realistic human situations requiring decision, however, alternatives often must be examined sequentially and decisions must be made as the alternatives are presented. When this is the case, and when the individual does not know what variables determine the order of the offers (and therefore does not know whether the future alternatives are likely to be "better" or "worse" than those with which he has already been confronted), he may well choose the first satisfactory alternative which is offered. In the terms of the present paper, the individual may choose the first alternative containing an outcome at or above his level of aspiration. The techniques suggested in this paper could be used to provide an empirical test of this hypothesis.

Simon offers another interesting hypothesis which might be tested in the context of the present approach. As the individual is being presented with a sequence of choice situations, he may find

that it is easy for him to discover satisfactory alternatives among them. If so, his level of aspiration will rise. On the other hand, he may find it difficult to discover satisfactory alternatives among those made available to him, and in this case his level of aspiration will fall. Simon suggests that such changes in level of aspiration would tend to guarantee the existence of satisfactory solutions to the decision situation, for failure to discover initially satisfying alternatives would depress the level of aspiration and thereby bring satisfactory solutions into existence.

#### SUMMARY AND CONCLUSIONS

This paper suggests that Lewinian theory concerning level of aspiration may be integrated with certain parts of decision theory. An achievement scale may be viewed as a scale of utility of the achievement goals. A formal definition of level of aspiration in terms of utility is offered. The problem of ascertaining a person's level of aspiration reduces to the problem of measuring his utility of the achievement goals. It is hypothesized that level of aspiration is associated with the largest distance on an individual's utility scale. If this is so, with a discrete set of goals, ordered metric measurement is sufficient for identifying a person's level of aspiration, since an ordered metric scale contains not only a ranking of the entities (achievement goals) but also a ranking of the distances between them. Certain experimental evidence which supports the hypothesis is summarized. Suggestions for future research are presented; these draw upon the ideas presented here, together with those of other workers in the fields of level of aspiration and decision theory.

In conclusion, it would seem that a useful behavioral model of decision making should include not only the concepts of utility and subjective probability, as

do the present models, but should also include a formulation of the effects of level of aspiration and reinforcement on utility. That is, the model should include recognition that utility has a model in its own right, in which the main concepts are level of aspiration ( $LA$ ) and reinforcement effects ( $R$ ).

In terms of such an extended model, it may be said that if various alternatives are available to an individual, he will choose from among these alternatives, toward each of which he has a subjective probability of attainment and a utility, so as to maximize subjectively expected utility,  $SEU$ . That is, the individual will choose so as to maximize

$$SEU = \sum_i p_i u_i,$$

where

$$u = f(LA, R).$$

#### REFERENCES

1. BECKER, S. W., & SIEGEL, S. Utility of grades: level of aspiration in a decision theory context. Unpublished paper, Pennsylvania State Univer., December, 1956.
2. BERNOULLI, D. Specimen theoriae novae de mensura sortis. *Comentarii Academiae Scientiarum Imperiales Petropolitanae*, 1738, 5, 175-192. (Trans. by L. Sommer in *Econometrica*, 1954, 22, 23-36.)
3. CHILD, I. L., & WHITING, J. W. M. Determinants of level of aspiration: evidence from everyday life. In H. Brand (Ed.), *The study of personality*. New York: Wiley, 1954. Pp. 495-508.
4. COOMBS, C. H. Psychological scaling without a unit of measurement. *Psychol. Rev.*, 1950, 57, 145-158.
5. COOMBS, C. H. A theory of psychological scaling. *Bull. Univer. Michigan Engng Res. Inst.*, 1952, No. 34.
6. COOMBS, C. H., & KOMORITA, S. S. Measuring utility of money through decision-making. Unpublished manuscript, Univer. of Michigan, February, 1955.
7. DAVIDSON, D., SIEGEL, S., & SUPPES, P. Some experiments and related theory in the measurement of utility and subjective probability. Stanford Value Project,



- Rep. No. 4, 1955. Also in DAVIDSON, D., SUPPES, P., & SIEGEL, S. *Decision-making: an experimental approach*. Stanford, Calif.: Stanford Univer. Press, 1957. Pp. 19-81.
8. DEMBO, TAMARA. Der Ärger als dynamisches Problem. *Psychol. Forsch.*, 1931, 15, 1-144.
9. EDWARDS, W. The theory of decision-making. *Psychol. Bull.*, 1954, 51, 380-417.
10. EDWARDS, W. The prediction of decisions among bets. *J. exp. Psychol.*, 1955, 50, 201-214.
11. ESCALONA, SYBILLE K. The effect of success and failure upon the level of aspiration and behavior in manic-depressive psychoses. *Univer. Ia. Stud. Child Welf.*, 1940, 16 (3), 199-302.
12. FESTINGER, L. A theoretical interpretation of shifts in level of aspiration. *Psychol. Rev.* 1942, 49, 235-250.
13. FRANK, J. D. Recent studies of the level of aspiration. *Psychol. Bull.*, 1941, 38, 218-226.
14. HOPPE, F. Erfolg and Misserfolg. *Psychol. Forsch.*, 1930, 44, 1-62.
15. HURST, P. M., & SIEGEL, S. Prediction of decisions from a higher ordered metric scale of utility. *J. exp. Psychol.*, 1956, 52, 138-144.
16. LEWIN, K., DEMBO, TAMARA, FESTINGER, L., & SEARS, PAULINE S. Level of aspiration. In J. McV. Hunt (Ed.), *Personality and the behavior disorders*. Vol. I. New York: Ronald, 1944. Pp. 333-378.
17. MOSTELLER, F., & NOGEE, P. An experimental measurement of utility. *J. polit. Econ.*, 1951, 59, 371-404.
18. RAMSEY, F. P. Truth and probability. In *The foundation of mathematics and other logical essays*. New York: Harcourt Brace, 1931, Chapter VII.
19. ROTTER, J. B. Level of aspiration as a method of studying personality: I. A critical review of methodology. *Psychol. Rev.*, 1942, 49, 463-474.
20. SAVAGE, L. J. *The foundations of statistics*. New York: Wiley, 1954.
21. SIEGEL, S. A method of obtaining an ordered metric scale. *Psychometrika*, 1956, 21, 207-216.
22. SIEGEL, S. *Nonparametric statistics*. New York: McGraw-Hill, 1956.
23. SIEGEL, S., & BELTZ, S. Empirical verification of a decision-theory model: prediction of choices between uncertain outcomes in the discrete case. Paper read at Eastern Psychol. Ass., Atlantic City, March, 1956.
24. SIEGEL, S., & SHEPHERD, IRMA. An ordered metric measure of social distance. Paper read at Eastern Psychol. Ass., Atlantic City, March, 1956.
25. SIMON, H. A. A behavioral model of rational choice. *Quart. Rev. Econ.*, 1955, 69, 99-118.
26. STEVENS, S. S. On the theory of scales of measurement. *Science*, 1946, 103, 677-680.
27. STEVENS, S. S. Mathematics, measurement, and psychophysics. In S. S. Stevens (Ed.), *Handbook of experimental psychology*. New York: Wiley, 1951. Pp. 1-49.
28. VON NEUMANN, J., & MORGENSTERN, O. *Theory of games and economic behavior*. (2nd ed.) Princeton: Princeton Univer. Press, 1947.

## A NOTE ON THE CRESPI EFFECT

A. C. PEREBOOM<sup>1</sup>

*Texas Technological College*

The classical Crespi (3) and the Tolman and Honzik (6) studies report abrupt shifts in performance following a discrete shift in the magnitude of the reward. The writer recently attempted to reconcile their data with his own where this "abruptness" effect failed to occur (4). This led to certain inferences about the role of exploratory activity in performance measures. In addition, Hull's habit construct was revised to improve its handling of the learning process following such a shift in reward magnitude. The present paper attempts to extend that study specifically to include Crespi's contrast effect.

It was argued in that paper that the early trials in most learning experiments involve much exploratory behavior which competes, performance-wise, with the one response measure that concerns *E*, that based on the acquisition of a goal expectancy. Learning about the maze or runway takes time, however, and knowledge about the goal forms only a small portion of the total learning. The poor performance in the early trials is a measure of this large amount of learning going on as it is reflected in exploratory behavior. Thus, even the runway does not give us a single response tendency (2). But as trials continue, exploration drops out and the competing goal response becomes more and more dominant: performance is now approaching an asymptote.

At this point in the Crespi experiment the reward magnitude was changed. All that his rats now had to learn was that a new magnitude was present; run-

way exploration, reflected in gradual improvement, was essentially completed during the initial learning. The response change was therefore abrupt relative to the original rate of response improvement.

Now consider the contrast effect. The level of runway familiarity or learning at the time of the reward-shift will be a function of the previous amount of runway exploration, and this in turn will be partially determined by the magnitude of the initial incentive. A large reward will compete more effectively with the exploratory tendency than a small one. This is inferred from the greater running speeds found under large rewards. Thus a large-reward group knows less about the runway than a small-reward group (cf. [1]). But unfamiliarity is a condition that arouses the exploratory tendency. If the large-reward group is now put on a smaller incentive they will not only begin "extinguishing" to some extent but they will also begin exploring—and to a greater degree than would a group on this small incentive throughout training, simply because they know less about this runway than would such a control group. The same reasoning should lead to the "elation" effect for the upward-shift group. An initially small reward means a large amount of initial exploration. Under a new and larger amount, they will need to explore less than would a group under that larger amount throughout training; they can therefore run faster than this group.

But the upward contrast effect does not always appear. Negative results were obtained in Spence's laboratory in several experiments designed along the lines of the Crespi study (5, pp. 127-

<sup>1</sup> The writer is indebted to Dr. Keith J. Hayes for his critical reading of this paper.

133). Spence concludes that the "elation" effect is due to a failure to train animals to their asymptotic speeds, additional trials being the source of the effect rather than the incentive shift. However, it may not be the asymptote per se. If the asymptotic speed is too close to some relative upper limit set by the given experimental situation, drive level, and quality of reward, there simply will not be enough room for the effect to show. To obtain the contrast effect then, the large reward must not be too large. Logically, of course, the same reasoning should apply to the downward effect. If the small reward leads to speeds little better than the operant speed, we would also fail to observe any effect.

A preliminary test of the present interpretation will introduce one condition into the Crespi study that was apparently omitted: Thoroughly familiarize the animals with the runway prior to Trial 1. This should lead to greater rates of response improvement under the initial incentives, and so tend to eliminate the relative abruptness of performance changes following the incentive shifts. In addition, the asymptotes of the original performance curves should be greater than those obtained by groups lacking preliminary exploration: An upward contrast effect would already be present. Whether positive and negative contrast effects will appear following the incentive shifts, however, depends upon

(a) the amount of preliminary exploration and (b) how long such groups are maintained under their initial amounts of reward. If no differences in runway familiarity are present at the time of such incentive shifts, then no contrast effects should follow.<sup>2</sup>

#### REFERENCES

1. BRUNER, J. S., MATTER, J., & PAPANEEK, M. L. Breadth of learning as a function of drive level and mechanization. *Psychol. Rev.*, 1955, 62, 1-10.
2. COTTON, J. W. Running time as a function of amount of food deprivation. *J. exp. Psychol.*, 1953, 46, 188-198.
3. CRESPI, L. P. Quantitative variation of incentive and performance in the white rat. *Amer. J. Psychol.*, 1942, 55, 467-517.
4. PEREBOOM, A. C. An analysis and revision of Hull's theorem 30. *J. exp. Psychol.*, 1957, 53, 234-238.
5. SPENCE, K. W. *Behavior theory and conditioning*. New Haven: Yale Univ. Press, 1956.
6. TOLMAN, E. C., & HONZIK, C. H. Introduction and removal of reward, and maze performance in rats. *Univ. Calif. Publ. Psychol.*, 1930, 4, 257-275.
7. ZEAMAN, D. Response latency as a function of the amount of reinforcement. *J. exp. Psychol.*, 1949, 39, 466-483.

(Received for early publication January 28, 1957)

<sup>2</sup> It might be noted here that Crespi's use of a particularly long runway (20 ft.) upon which speed was recorded, rather than just latency, no doubt favored the exploratory tendency and thus gave him clear-cut results (cf. [7]).



# THE PSYCHOLOGICAL REVIEW

## THE PSYCHOLOGICAL SIGNIFICANCE OF THE CONCEPT OF "AROUSAL" OR "ACTIVATION"

ELIZABETH DUFFY

*The Woman's College of the University of North Carolina*

The concept of "arousal," "activation," or "energy mobilization," as developed by the writer over a period of many years (7, 9, 10, 11, 13), and employed by others in various contexts (15, 18, 25, 40), has wide applicability in psychology.<sup>1</sup> A fuller discussion of the topic will be presented elsewhere. Pending its appearance, however, it may be of interest to point out some of the areas which this concept should serve to illuminate.

It has been argued in previous papers (10, 12) that all variations in behavior may be described as variations in either the direction<sup>2</sup> of behavior or the intensity of behavior. Only one part of this argument is essential for the present purpose. Whatever may be the reaction to the attempt to reduce the descriptive categories of psychology to two

basic types of concept, we can proceed without dispute provided only it is agreed that intensity is a characteristic of behavior which can be abstracted and studied separately. It is the intensity aspect of behavior which has been variously referred to as the degree of excitation, arousal, activation, or energy mobilization.

I have argued that such abstraction from the totality of behavior is a necessary procedure if the psychologist is to be enabled to manipulate variables in a way likely to provide solutions to some of his problems. Confusion of the direction of behavior with the intensity of behavior, resulting in their fortuitous combination in certain psychological concepts (10) and in the "trait" names used to describe personality (12), was suggested as a possible basis for some of the unrewarding findings in many psychological investigations. Since the intensity of response can vary independently of the direction of response, it was proposed that it should be measured independently and its correlates investigated.

Perhaps a parallel may be seen in the analysis of sensory function.<sup>3</sup> Before

<sup>3</sup> For this suggestion of a parallel, I am indebted to Dr. R. B. Malmö, who, in the fall of 1955, was kind enough to read the major portion of my manuscript for a forthcoming book, and to discuss it with his staff.

<sup>1</sup> The terms "activation" and "arousal," as used here, do not refer specifically to the activation pattern in the EEG. On the contrary, they refer to variations in the arousal of the individual as a whole, as or excitation of the individual as a whole, as indicated roughly by any one of a number of physiological measures (e.g., skin resistance, muscle tension, EEG, cardiovascular measures, and others). The degree of arousal appears to be best indicated by a combination of measures.

<sup>2</sup> "Direction" in behavior refers merely to the fact that the individual does "this" rather than "that," or responds positively to certain cues and negatively to others.

progress could be made in the study of sensation and its physical correlates, it was necessary to separate the dimension of intensity from that of other sensory characteristics. In audition, for example, loudness was distinguished from pitch, and was related to a different type of variation in the physical stimulus. In vision, brightness was separated from hue, and each of these aspects of vision was related to the appropriate type of variation in the stimulus. Little progress in the understanding of sensation could have been made until suitable abstractions from the total sensory experience had been achieved, and these identifiable aspects of the totality had been investigated separately.

Measurement of the intensity of response (i.e., the degree of excitation, arousal, activation, or energy mobilization), it has been pointed out, may be achieved, at least in rough fashion, through various means (9, 10, 13, 15). Among the physiological measures which may be employed are skin conductance, muscle tension, the electroencephalogram (EEG), pulse rate, respiration, and others. These measures show intercorrelations, although the correlation coefficients are not always high since there is patterning in the excitation of the individual, the nature of which appears to depend upon the specific stimulus situation and upon organic factors within the individual.<sup>4</sup> Nevertheless, there is evidence also of "generality" of the excitation. Hence a concept of arousal, or energy mobilization, appears to be justified.

It should be noted that the physiological measures which serve as indicants of arousal, and which correlate at least to some degree with each other, include

measures of autonomic functions, of skeletal-muscle functioning, and of the functioning of the higher nerve centers. It is clear that it is the *organism*, and not a single system, or a single aspect of response, which shows arousal or activation.

The historical roots of the concept of activation lie in Cannon's concept of "energy mobilization" during "emotion" (3). Unlike Cannon's concept, however, the present concept of activation or arousal is designed to describe the intensity aspect of *all* behavior (10, 12). Referred to as the "degree of excitation," it was, in 1934, defined as "the extent to which the organism as a whole is activated or aroused" (9, p. 194). Both its definition and its proposed mode of measurement have in more recent publications followed the line suggested at that time (10, 13). When, however, studies of the electroencephalogram provided data on the behavioral correlates of changes in the EEG, it was suggested that this measure also provided an indication of the degree of arousal (13).

To those unfamiliar with the concept of activation, confusion frequently arises between the degree of internal arousal (referred to by the concept) and the vigor and extent of overt responses. While the degree of internal arousal usually correlates fairly closely with the intensity of overt response, a discrepancy between the two may be introduced by the intervention of inhibitory processes, a phenomenon which has not received the degree of attention to which it is entitled. An additional source of confusion is the tendency on the part of some to confuse activation or excitability with vitality. Actually, it is suggested that these two characteristics are more likely to be negatively related than to be positively related. The tendency to be frequently and intensely aroused

<sup>4</sup> The patterning of excitation is discussed more fully in the manuscript referred to in Footnote 3. It is believed that a more adequate concept of excitation, or activation, is thereby developed.

leads no doubt to fatigue and to a consequent reduction in vitality.

The chief point in regard to arousal, which I have repeatedly made (10, 11, 12, 13), is that arousal occurs in a *continuum*, from a low point during deep sleep to a high point during extreme effort or great excitement, with no distinguishable break for such conditions as sleep or "emotion." Evidence supporting this contention has been presented specifically for skin conductance, muscle tension, and the EEG (13). Recently Lindsley has elaborated upon the conception as it applies to the EEG (25), although earlier, in his "activation theory of emotion" (24, pp. 504-509), he had been of the opinion that "emotion" and sleep were conditions which were correlated with certain changes in the EEG, while conditions intermediate between the two were held to be as yet unexplained.

The factors which produce variations in the degree of arousal are various. They include, apparently, drugs, hormones, variations in physical exertion, and variations in what is commonly referred to as the degree of motivation. It appears that differences in the degree of arousal in different individuals may have a genetic or an environmental basis, or both. This conclusion is suggested from animal studies and from the relatively few studies of human beings in which the problem has been considered.

One of the potential contributions to psychology of the concept of arousal is that of breaking down the distinction between "drives" or "motives" and "emotion" (10, 11). The same kinds of physiological changes may be observed to occur in these variously designated conditions, and, depending upon the degree of arousal, to produce the same sorts of effect upon behavior. It has been contended that "emotion" is in no sense a unique condition, and that our

investigations should not be directed toward the study of "emotion" as such (9).

In the study of "motivation," the concept of arousal is of distinct service. By means of the physiological measures which serve as indicants of arousal, we may secure a direct measure of the degree (intensity) of "motivation."<sup>5</sup> Any other measure must of necessity be less direct. When all factors affecting the level of arousal except the degree of incentive value or threat value are held constant, measurement of the degree of arousal affords a measure of the "motivating" value of a given situation. It also affords, incidentally, an objective measure of what is called the "stress" imposed by a situation.

Physiological measurements made in a wide variety of situations have shown the expected correspondence between the degree of arousal and the apparent degree of significance of the situation—i.e., its incentive value or its threat value (13). For example, men undergoing flight training were found to show more tension of the muscles during the solo stage of training than during other stages, and during the maneuvers of take-off and landing than during other maneuvers (39). Galvanic skin responses obtained during replies to questions about provocative social problems were found to be smaller if the replies were in harmony with group opinion than if they were not, and "Yes" responses were found in general to be associated with smaller galvanic reactions than "No" responses (34).

The concept of activation holds fur-

<sup>5</sup> The concept of "motivation," as currently employed, is a "compound" concept which incorporates a description of both the "drive level," or arousal aspect, of behavior and also the direction taken by behavior, i.e., the selectivity of response. These two aspects of behavior may vary independently, though both are characteristically affected by a certain stimulus-condition such as hunger.



ther significance for psychology by virtue of the fact that variations in the degree of activation are, on the average, accompanied by certain variations in overt response.<sup>6</sup> The degree of activation appears to affect the speed, the intensity, and the coordination of responses. In general, the optimal degree of activation appears to be a moderate one, the curve which expresses the relationship between activation and quality of performance taking the form of an inverted U. This conclusion, as it relates to muscular tension and performance, was suggested by me in 1932 (8, pp. 544-546), by Freeman in several papers published around that time (15), and later by Courts (4). That it holds also for other indicators of the degree of activation is suggested by Freeman's finding that skin resistance and reaction time, measured simultaneously on a single subject for 105 trials over a number of days, gave an inverted U-shaped curve when plotted on a graph (14). More recently the EEG has been found to show the same sort of relationship to reaction time (22).

The effect of any given degree of activation upon performance appears to vary, however, with a number of factors, including the nature of the task to be performed and certain characteristics of the individual—such as, perhaps, the ability to inhibit and coordinate responses under a high degree of excitation (8). Organismic interaction is the basic explanatory principle suggested to account for the particular effects upon performance of various degrees of activation. Such organismic interaction may also, it appears, have some effect upon sensory thresholds. Again the possibility presents itself that the relationship may take the form of an inverted U-shaped curve.

<sup>6</sup> These studies are reviewed in the manuscript referred to in Footnote 3.

When performance has been observed to vary under certain conditions, such as those of drowsiness, of fatigue, or of "emotion," it is suggested that the variation may be due, at least in part, to the effect of varying degrees of arousal. The disorganization of responses frequently reported during "overmotivation" or "emotion," for example, may be conceived of as resulting in part from too high a degree of arousal. Such a condition would be represented at one end of the U-shaped curve. A similar disorganization of responses, found sometimes during drowsiness or fatigue, would be represented at the other end of the curve showing the relationship between arousal and performance. In any case, it seems clear that prediction of overt response to a given set of stimulating conditions can be increased in accuracy when there is knowledge of the degree of internal arousal.

It appears also that, under similar stimulation, individuals differ in the degree of their arousal and in the speed with which they return to their former level of functioning. Moreover, there is evidence of consistency in this individual variation. Apparently the individual who responds with intensity in one situation will, on the average, respond with intensity in other situations also, as compared with other individuals. While the degree of arousal varies with the situation, the rank in arousal tends to be preserved. Different individuals appear to vary around different central tendencies—i.e., to differ in responsiveness. The easily aroused, or more responsive, individual has been found to show this responsiveness in many different forms, some of which will be described below.

For instance, subjects who showed a large number of galvanic skin responses when there was no observable stimulation also showed less adaptation of the galvanic skin response (GSR) to repeated stimulation (33).

Similarly, the frequency of the alpha rhythm in normal adults has been reported to show a significant relationship to ratings on the behavioral continuum called "primary-secondary functioning" (32). Individuals in whom the alpha rhythm was more rapid tended to show more "primary functioning," or to be "quick, impulsive, variable, and highly stimutable." Those with relatively low frequencies of the alpha rhythm tended to show more "secondary functioning," or to be "slow, cautious, steady, with an even mood and psychic tempo. . . ." Mundy-Castle hypothetically ascribed these behavioral differences to differences in excitability within the central nervous system, the "primary functioning" individuals showing the greater excitability. A difference in neural excitability was also suggested as the explanation of his finding that there was a significant difference in the EEG activity evoked by rhythmic photic stimulation between subjects with a mean alpha frequency above 10.3 cycles per second and those with a mean alpha frequency below that rate.<sup>7</sup> He offered the same explanation of the greater incidence of "following"<sup>8</sup> in the beta range by those subjects showing little alpha rhythm, even when the eyes were closed, as compared with those subjects showing persistent alpha rhythms (32).

Gastaut and his collaborators have also reported individual differences in cortical excitability (17). While their major purpose was not the investigation

of individual differences, they made the incidental observation that calm individuals had a slow, high-voltage alpha rhythm (8-10 c./s.), with little "driving" of occipital rhythms by photic stimulation. Neurons showed a long recuperation time, synchrony of response was said to be noticeable, and recruitment poor. "Nervous" individuals, on the other hand, were said to have a high-frequency, low-voltage alpha rhythm (10-13 c./s.), which at times was not perceptible. They were described as having a short neuronal recuperation time, little synchrony of response, good recruitment, and considerable driving by photic stimulation. In other words, "calm" as compared with "nervous" individuals showed less cortical excitability.

Differences in the EEG's of different individuals under similar stimulating conditions appear to be correlated also with differences in another form of responsiveness—i.e., differences in the threshold of deep reflexes. It has been reported that normal subjects with deep reflexes which are difficult to elicit showed a high percentage of alpha activity and little or no fast activity, while those with deep reflexes which were hyperactive had little alpha activity and a high percentage of fast activity (21). However, while groups at the two extremes of reflex responsiveness differed significantly in the percentage of alpha activity, there was wide variation in the extent of such activity within any one of the groups formed on the basis of reflex status. Amplitude of rhythm was observed to be greatest in EEG records showing pronounced alpha activity.

Proneness to develop anxiety under stress, which may perhaps be regarded as a form of hyperresponsiveness, has been found, in both normal subjects and psychiatric patients, to be associated with a significantly smaller percentage

<sup>7</sup> It is believed, he says, that "electrical rhythms in the brain can be initiated or augmented by a process similar to resonance; in other words, if an area of the brain is subjected to rhythmic impulses corresponding to its own latent or actual frequency, it may itself oscillate for as long as stimulation is maintained" (33, p. 319). It is thought that the area may also be activated by stimulation harmonically related to its own.

<sup>8</sup> "Following" refers to electrical responses in the cortex occurring at the stimulus frequency.



of resting brain-wave activity in the alpha region when this activity is determined by automatic frequency analysis (35). The anxiety-prone groups showed more fast activity (16-24 c./s.), or more slow activity (3-7 c./s.), below the alpha range. The significance of the slow activity is not as clear as that of the fast activity. Fast activity may be presumed to be indicative of a high level of excitation. It has been observed, for example, at the beginning of EEG recording in normal subjects who are unusually apprehensive about the procedure, and it has been found to disappear with reassurance and the attainment of relaxation (24). It appears at least possible that the slow activity may be due to fatigue from previous states of intense arousal.

In an investigation employing prison farm inmates, schizophrenics, and control subjects, to whom a group of psychological tests were given, it was reported that EEG activity above 16-20 c./s. appeared in significant amounts only in the records of those who, as rated by the psychological tests, showed anxiety to a marked degree (20). Slow activity was said not to be very prevalent, but when it did occur, to be found most often among the patients.

These and other studies suggest that anxiety-proneness may be conceived of as a form of overarousal or hyperresponsiveness. The EEG's of the anxiety-prone seem very similar in most instances to the EEG's of other subjects whose exceptional responsiveness to the environment is indicated by active reflexes, or by ratings on "primary function."

Degree of tension of the skeletal muscles is another indicator of responsiveness, or ease and extent of arousal, in which differences between individuals have been found. In almost every investigation in which tension of the skeletal musculature has been measured,

wide differences between individuals in the degree of tension have been noted.<sup>9</sup> In the same stimulus situation, one individual would respond with a relatively low degree of tension, another with a moderate degree, and a third with a high degree of tension. Moreover, when observed in a *different* stimulus situation, the subjects, while varying in their absolute level of tension, would tend to preserve their ranks with respect to tension of the muscles. It was thus shown that different individuals vary around different central tendencies, so that one individual might be characterized as being in general tense, and another as being in general relaxed.

In early studies of muscular tension, the writer found, in two separate investigations, that nursery school children showed marked individual differences in grip pressure while engaged in various tasks, and that there was a significant correlation between the grip pressure on one occasion and that on another, and during one task and during another (6, 7). Grip pressure scores were found to be independent of the strength of grip as indicated by dynamometer scores, but to be related to ratings on excitability and on adjustment to the nursery school, the tense children being rated as more excitable and, on the average, less well adjusted.

Arnold also found that individuals tended to preserve their rank in the group with respect to pressure from the hand during repetition of the same task and during the performance of different tasks (2).

A study of airplane pilots in training revealed that some showed excessive muscle tension (pressure on stick and on rudder pedal) in both take-offs and

<sup>9</sup> Differences in muscle tension will, for the purposes of this discussion, refer to differences in pressure exerted by some group of muscles or to differences in electric potentials from muscles.



landings, while others showed little tension on either maneuver (39). No individuals were found who in general tended to be tense during take-offs alone or during landings alone.

Further evidence that individuals who are more highly activated than others in one stimulus situation, as indicated by tension of the skeletal muscles, are more responsive to a wide variety of stimuli, is presented in studies by Lundervold (26). "Tense" subjects, as compared with "relaxed" subjects, were found to show more activity in the muscles when external conditions were changed, as by an increase in noise, the lowering of the room temperature, or the introduction of certain stimuli which caused irritation or anger. In these persons, there was not only more activity in the single muscle, but also electrical activity in more muscles, including muscles which did not participate directly in the movement. At the end of thirty minutes of noise, fifty per cent of the tense subjects, as compared with none of the relaxed subjects, showed more action potentials than they had shown before the noise began.

A similar relationship between muscular tension and another form of responsiveness was earlier shown by Freeman and Katzoff, who found a significant correlation between grip-pressure scores and scores on the Cason Common Annoyance Test (16). Subjects with higher pressure scores tended to be more frequently or intensely annoyed—i.e., to show greater responsiveness of the sort referred to as "irritability."

It appears that, on the whole, skeletal-muscle tension in one part of the body tends to be positively related to that in other parts of the body, though the relationship between the tension in any two areas may not be very close. Parts of the body more remote from each other, or more widely differentiated in function, yield tension measures which

are less closely related than those which are closer together or functionally more similar. When tension measures taken from different parts of the body, recorded during different tasks, or made at widely separated intervals of time, nevertheless show a significant positive correlation with each other, it must, however, be concluded that there is at least some degree of "generality" in skeletal-muscle tension. Moreover, from measuring the responsiveness of the skeletal-muscle system, we may apparently predict to some extent the response of highly integrated systems of reaction described as "personality traits." Indeed, in a study in which no direct measure of muscular tension was employed, but in which ratings on muscular tension and measures of sixteen physiological variables were intercorrelated and submitted to factor analysis, a factor defined as muscular tension showed correlation with certain personality characteristics (36).

Since conditions of high activation may perhaps increase the likelihood of disorganization of motor responses, it is not surprising that measures of tremor and other forms of motor disorganization have been found to be related to the severity of conflicts (31) and to neuroticism (1, 19, 23, 28, 29, 30). Measures of irregularity in pressure appear to be among the measures which discriminate best between a normal and a psychiatric population, a finding which might be expected if, as suggested by the writer (8) and by Luria (27), irregular pressure tracings are indicative of poor coordination or lack of control of responses.

Other indicants of arousal have also been shown to be related to more complex forms of response. For example, it has been said that a reasonably accurate prediction of a person's respiratory rate at a given time during a flight could be made on the basis of knowl-

edge of his "normal" respiratory rate and the name of the maneuver to be performed (39).

Similarly, when an "autonomic factor" was obtained from twenty physiological measures related to the functioning of the autonomic nervous system, it was found that individuals differed greatly in scores on this factor, but that the correlation coefficient between early and later factor scores did not drop below .64 over a two-year period (38). Children at one extreme of the autonomic-factor scores were reported to differ significantly from those at the other extreme in certain personality traits (37).

Individuals differ, not only in the degree of excitation produced by stimulation, but also in the speed with which the processes affected return to their prior level of functioning. Moreover, differences in recovery time cannot be accounted for solely by differences in the degree of arousal, for they are found when recovery is measured *in relation to the degree of arousal*. Darrow and Heath, who first made use of this measure, computed a "recovery-reaction quotient" by dividing the extent of recovery in skin resistance by the extent of decrease in resistance which had occurred as a result of stimulation (5). The recovery-reaction quotient was reported to be related to many different measures of "neurotic" and emotionally unstable tendencies." The investigators concluded that it was one of their best indicators of the absence of neurotic trend, but that the coefficients of correlation were not high enough to justify the use of the measure for prediction in individual cases. It would appear that the speed of recovery from arousal is an extremely significant aspect of response, and one which deserves further investigation.

Individuals who are exceptionally responsive to the environment may show their responsivity in behavior which,

from a directional point of view, may be described in diverse ways. A tendency toward a high degree of arousal does not determine which aspects of the environment an individual will approach or will have a tendency to approach (i.e., have a favorable attitude toward); nor does it determine which aspects of the environment he will withdraw from or have a tendency to withdraw from (i.e., have an unfavorable attitude toward). On the contrary, the orientation of the individual in his environment is determined largely by other factors. These are, of course, the factors, both genetic and environmental, which have given to various aspects of his environment the nature of their significance, or their "cue-function." There are, nevertheless, differences in the way in which approach or withdrawal occurs which may conceivably be derived from differences in the level of activation. Among these appear to be differences in such aspects of behavior as alertness, impulsiveness, irritability, distractibility, and the degree of organization of responses. Moreover, greater responsiveness may, it is suggested, facilitate the development of aggression or withdrawal, enthusiasm, or anxiety. The more responsive individual in a certain kind of environment is no doubt more susceptible to the effects of that environment. Presumably he may become, depending upon circumstances, more anxiety-prone, more conscientious, more sympathetic, more devoted, or more irascible than a less responsive person would become under similar circumstances. We should therefore expect to find some association between a high degree of activation and easily aroused or intense responses of various kinds (e.g., anxieties, resentments, enthusiasms, or attachments). From knowledge of the individual's tendencies with respect to activation we should not, however, be able to predict the direction which his behavior



would take. A more dependable association might be expected between individual differences in excitability and differences in the "dynamic" characteristics of behavior such as those mentioned above.

The effect of a high degree of arousal upon overt behavior varies, no doubt, with variations in the degree of inhibitory ability (9), or, as Luria has described it, with variations in the strength of the "functional barrier" between excitation and response (27).<sup>10</sup> Depending upon this factor, a high degree of activation may, I suggest, lead to impulsive, disorganized behavior or to sensitive, alert, vigorous, and coordinated responses to the environment. Evidence in support of these statements is at present so meager, however, as to leave them in the category of speculations. It is to be hoped that further investigation will provide the basis for a more confident statement of the relationship between "personality" characteristics and individual differences in the level of activation.

#### SUMMARY

The concept of arousal or activation appears to be a significant one for the ordering of psychological data. Differences in activation, as shown in a wide variety of physiological measures, appear to be associated with many other differences in response.

In different stimulus-situations, the same individual differs in the degree of arousal. Measurement of the physiological indicants of arousal affords, when

other factors are constant, a direct measure of the "motivating" or "emotional" value of the situation to the individual. The concept serves to break down the distinction between the arousal aspect of "drives" or "motives" and that of "emotion," and to suggest instead a continuum in the degree of activation of the individual.

Differences in activation in the same individual are, it is suggested, accompanied by differences in the quality of performance; the relationship may be graphically represented by an inverted U-shaped curve. Further data are needed, however, to establish the validity of this hypothesis.

In the same stimulus situation there are differences between individuals in the degree of arousal. These differences tend to persist, and thus to characterize the individual. Moreover, the easily aroused, or responsive, person shows this responsiveness in many forms. It has been observed in the ease with which deep reflexes are elicited, and in the extent, frequency, and duration of reactions to stimulation, both of the skeletal musculature and of various functions controlled by the autonomic nervous system. It has been shown also in differences in cortical potentials, which are presumably indicative of differences in the excitability of higher nerve centers. These various forms of responsiveness show, in general, positive intercorrelations, though the coefficients of correlation are apparently not high enough for a measure of any one mode of responsiveness to serve as an adequate measure of the general responsiveness of the individual. They appear, however, to give justification to the conception of a responsive or an unresponsive *individual*, not merely responsive or unresponsive skeletal musculature, skin resistance, or cortical potentials.

Differences in arousal are shown also in responses of greater inclusiveness and

<sup>10</sup> Luria reports that children show weakness of the functional barrier between excitation and motor response, as indicated by poor performance on a test requiring that a key be pressed down as slowly as possible (28). The writer noted that, during a discrimination performance, younger nursery school children, with irregular grip-pressure tracings, had a higher proportion of their errors in the category of "impulsive" errors, or errors of over-reaction (8).



of higher integration—i.e., in responses frequently classified as personality traits. Combining with one or another directional aspect of behavior, a persistent high degree of arousal may, it appears, be observed in many complex characteristics, such as anxiety-proneness or aggressiveness.

Facts such as those presented above suggest that the concept of activation may prove useful in many different areas of psychology.

#### REFERENCES

1. ALBINO, R. C. The stable and labile personality types of Luria in clinically normal individuals. *Brit. J. Psychol.*, 1948, 39, 54-60.
2. ARNOLD, M. B. A study of tension in relation to breakdown. *J. gen. Psychol.*, 1942, 26, 315-346.
3. CANNON, W. B. *Bodily changes in pain, hunger, fear and rage*. New York: Appleton, 1915, 1929.
4. COURTS, F. A. Relations between muscular tension and performance. *Psychol. Bull.*, 1942, 39, 347-367.
5. DARROW, C. W., & HEATH, L. L. Reaction tendencies related to personality. In K. S. Lashley (Ed.), *Studies in the dynamics of behavior*. Chicago: Univer. of Chicago Press, 1932. Pp. 59-261.
6. DUFFY, E. Tensions and emotional factors in reaction. *Genet. Psychol. Monogr.*, 1930, 7, 1-79.
7. DUFFY, E. The measurement of muscular tension as a technique for the study of emotional tendencies. *Amer. J. Psychol.*, 1932, 44, 146-162.
8. DUFFY, E. The relationship between muscular tension and quality of performance. *Amer. J. Psychol.*, 1932, 44, 535-546.
9. DUFFY, E. Emotion: an example of the need for reorientation in psychology. *Psychol. Rev.*, 1934, 41, 184-198.
10. DUFFY, E. The conceptual categories of psychology: a suggestion for revision. *Psychol. Rev.*, 1941, 48, 177-203.
11. DUFFY, E. An explanation of "emotional" phenomena without the use of the concept "emotion." *J. gen. Psychol.*, 1941, 25, 283-293.
12. DUFFY, E. A systematic framework for the description of personality. *J. abnorm. soc. Psychol.*, 1949, 44, 175-190.
13. DUFFY, E. The concept of energy mobilization. *Psychol. Rev.*, 1951, 58, 30-40.
14. FREEMAN, G. L. The relationship between performance level and bodily activity level. *J. exp. Psychol.*, 1940, 26, 602-608.
15. FREEMAN, G. L. *The energetics of human behavior*. Ithaca: Cornell Univer. Press, 1948.
16. FREEMAN, G. L., & KATZOFF, E. T. Muscular tension and irritability. *Amer. J. Psychol.*, 1932, 44, 789-792.
17. GASTAUT, H. ET Y., ROGER, A., CORRIOL, J., & NAQUET, R. Étude électrographique du cycle d'excitabilité cortical. *EEG clin. Neurophysiol.*, 1951, 3, 401-428.
18. HEBB, D. O. Drives and the C.N.S. (conceptual nervous system). *Psychol. Rev.*, 1955, 62, 243-254.
19. JOST, H. Some physiological changes during frustration. *Child Developm.*, 1941, 12, 9-15.
20. KENNARD, M. A., RABINOVITCH, M. S., & FISTER, W. P. The use of frequency analysis in the interpretation of the EEG's of patients with psychological disorders. *EEG clin. Neurophysiol.*, 1955, 7, 29-38.
21. KENNARD, M. A., & WILLNER, M. D. Correlation between electroencephalograms and deep reflexes in normal adults. *Dis. nerv. System*, 1943, 6, 337-347.
22. LANSING, R. W., SCHWARTZ, E., & LINDSLEY, D. B. Reaction time and EEG activation. *Amer. Psychologist*, 1956, 11, 433.
23. LEE, M. A. M. The relation of the knee jerk and standing steadiness to nervous instability. *J. abnorm. soc. Psychol.*, 1931, 26, 212-228.
24. LINDSLEY, D. B. Emotion. In S. S. Stevens (Ed.), *Handbook of experimental psychology*. New York: Wiley, 1951. Pp. 473-516.
25. LINDSLEY, D. B. Psychological phenomena and the electroencephalogram. *EEG clin. Neurophysiol.*, 1952, 4, 443-456.
26. LUNDVOLD, A. An electromyographic investigation of tense and relaxed subjects. *J. nerv. ment. Dis.*, 1952, 115, 512-525.
27. LURIA, A. R. *The nature of human conflict* (Transl. and ed. by W. H. Gantt). New York: Liveright, 1932.
28. MALMO, R. B., SHAGASS, C., BÉLANGER, D. J., & SMITH, A. A. Motor control in psychiatric patients under experimen-

- tal stress. *J. abnorm. soc. Psychol.*, 1951, 46, 539-547.
29. MALMO, R. B., SHAGASS, C., & DAVIS, J. F. Electromyographic studies of muscular tension in psychiatric patients under stress. *J. clin. exp. Psychopath.*, 1951, 12, 45-66.
  30. MALMO, R. B., & SMITH, A. A. Forehead tension and motor irregularities in psychoneurotic patients under stress. *J. Pers.*, 1955, 23, 391-406.
  31. MORGAN, M. I., & OJEMANN, R. H. A study of the Luria method. *J. appl. Psychol.*, 1942, 26, 168-179.
  32. MUNDY-CASTLE, A. C. Electrical responses of the brain in relation to behavior. *Brit. J. Psychol.*, 1953, 44, 318-329.
  33. MUNDY-CASTLE, A. C., & MCKIEVER, B. L. The psychophysiological significance of the galvanic skin response. *J. exp. Psychol.*, 1953, 46, 15-24.
  34. MURRAY, H. A. *Explorations in personality*. New York: Oxford Univer. Press, 1938.
  35. ULETT, G. A., GLESER, G., WINOKUR, G., & LAWLER, A. The EEG and reaction to photic stimulation as an index of anxiety-proneness. *EEG clin. Neurophysiol.*, 1953, 5, 23-32.
  36. WENGER, M. A. An attempt to appraise individual differences in level of muscular tension. *J. exp. Psychol.*, 1943, 32, 213-225.
  37. WENGER, M. A. Preliminary study of the significance of measures of autonomic balance. *Psychosom. Med.*, 1947, 9, 301-309.
  38. WENGER, M. A., & ELLINGTON, M. The measurement of autonomic balance in children: method and normative data. *Psychosom. Med.*, 1943, 5, 241-253.
  39. WILLIAMS, A. C., JR., MACMILLAN, J. W., & JENKINS, J. G. *Preliminary experimental investigations of "tension" as a determinant of performance in flight training*. Civil Aeronautics Admin., Div. of Res., Rep. No. 54, Washington, D. C. January, 1946.
  40. WOODWORTH, R. S., & SCHLOSBERG, H. *Experimental psychology*. (Rev. ed.) New York: Holt, 1954.

(Received October 22, 1956)

## ANXIETY AND BEHAVIORAL AROUSAL<sup>1</sup>

ROBERT B. MALMO<sup>2</sup>

*Allan Memorial Institute of Psychiatry, McGill University*

During the past two decades there has been a growing interest in objective physiological studies of psychiatric patients. In this work, one of the most prominent psychological concepts has been that of anxiety. Although there is general agreement that the areas denoted by the term "anxiety" are important ones for study, there is nonetheless considerable disagreement concerning what the term means. In large measure, this semantic difficulty is part of a larger problem facing psychology today, and that is to find a way out of the confusion surrounding the concepts of motivation and emotion. Duffy has cogently argued that these concepts are second-order ones which reduce to primary factors of intensity and direction, and that along the intensity dimension, at least, the distinction between motivation and emotion is unnecessary (9, 10, 11).<sup>3</sup>

This is not to say that the directional aspect is not important or to deny that,

in terms of direction, meaningful distinctions may be made between motivation and emotion, and indeed between different emotions. However, for present purposes it is essential to focus on the question of what these phenomena have in common rather than to consider their differences; in this paper, therefore, we shall be primarily concerned with the intensity dimension.

The main purpose of the present paper is to consider recent experimental data in an attempt to find a way out of the present confusion. I shall begin with a summary of two lines of investigation in our laboratory, dealing first with our discovery that certain physiological measures may serve as indicants of intensity or "behavioral arousal." These experiments were performed with nonpatient subjects. Second, in summarizing our investigations of pathological anxiety in psychiatric patients, I shall attempt to use the concept of behavioral arousal in an integrative way. Third, I shall draw on data from recent neurophysiological investigations to indicate possible mechanisms involved in the pathology and etiology of anxiety. Finally, on the basis of these theoretical considerations, I suggest problems requiring further experimental study.

### PHYSIOLOGICAL INDICANTS OF BEHAVIORAL INTENSITY

In 1951 we (31) reported finding a gradient phenomenon from electromyographic (EMG) recording during mirror tracing. Since that time the phenomenon has been observed under various conditions in our laboratory. Figure 1 presents mirror-drawing data from a study by Bartoshuk (1). Note that the

<sup>1</sup> This paper reviews work which was supported by the Medical Research and Development Division, Office of the Surgeon General, Department of the U. S. Army, under Contract Number DA 49-007-MD-626, by Defence Research Board Grant Number 9425-04 (Canada), and by Grant Number A.P. 29 from the National Research Council of Canada.

<sup>2</sup> The author is indebted to Drs. A. K. Bartoshuk, D. Bindra, F. R. Brush, D. E. Cameron, D. O. Hebb, and R. G. Stennett for criticizing earlier drafts of this paper.

<sup>3</sup> I do not wish to imply that this has been Duffy's only theoretical contribution. Her writings contain prior reference to a dimension of behavioral intensity (conceived as a continuum of "arousal," or "activation"); and she has previously cited evidence to support the argument that physiological measures may serve as the chief means of quantifying such a dimension or continuum.



chin lead (which taps the speech muscles) also shows a gradient—that is, progressively rising muscle potentials from the beginning to the end of the task. Bélanger (3) found similar gradients from the arm in a size-discrimination task. Wallerstein (42) found gradients in the frontalis muscle in a task about as completely devoid of motor components as one could possibly design. The subject, reclining on a comfortable bed, listened to verbal material (short detective story or essay) presented to him by a tape recorder. In Wallerstein's experiment, the gradients extended over ten minutes and their steepness was related to the subject's reported degree of interest in listening (2, p. 228 f.).

Bartoshuk (2) was the first to show that the fastest and most accurate subjects (i.e., superior performers on mirror tracing) produced the steepest muscle-potential gradients. Such a relationship of EMG gradients to motivation has

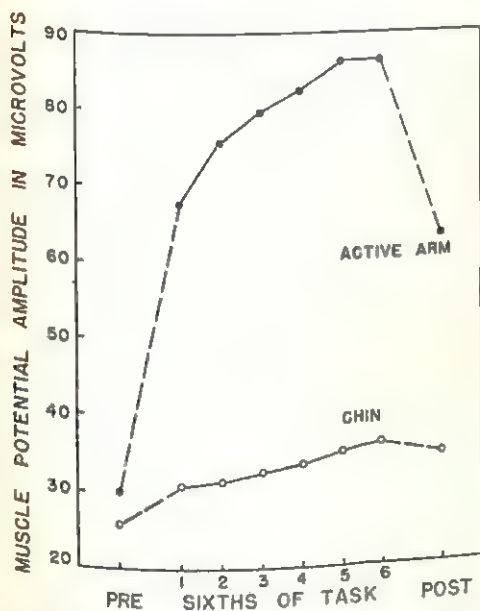


FIG. 1. Graphs showing mean EMG gradients in Bartoshuk's experiment (1). Note that gradient was also obtained from chin lead which records from muscles of speech.  $N = 17$ .

been confirmed by three subsequent studies, employing tracking tasks. Surwillo (39) demonstrated that raising incentive had the effect of increasing the steepness of EMG gradients in a visual tracking experiment. Figure 2 presents confirmatory data from a more recent experiment by Stennett (37) who employed auditory tracking under four conditions, with increasing degrees of incentive. Note that the muscle potentials were recorded from the nonactive, left arm. His "exertion" condition merely involved the subject's holding the tracking knob over at a fixed point in order to control for sheer physical work. Under the "calibration" condition the subject believed that he was just assisting with calibration of the apparatus, and that his tracking scores were not being recorded. The "optimal" condition was designed to motivate the subject sufficiently to elicit his most efficient performance, whereas the "incentive" condition was designed to "overmotivate" the subject by offering large bonuses for high-level performance and threatening with strong electric shock if performance did not reach this high level. The differences shown in the figure were statistically significant. In brief, Stennett's findings indicated that the most efficient tracking performance was associated with intermediate physiological levels (i.e., intermediate steepness of EMG gradients and intermediate levels of palmar skin conductance). With lower levels of physiological functioning (less steep gradients, lower levels of palmar skin conductance), performance on tracking was inferior. However, going now to the other extreme, performance on tracking associated with extremely high EMG gradients and extremely high palmar skin conductance was also inferior to tracking performance associated with moderately high levels of physiological functioning.

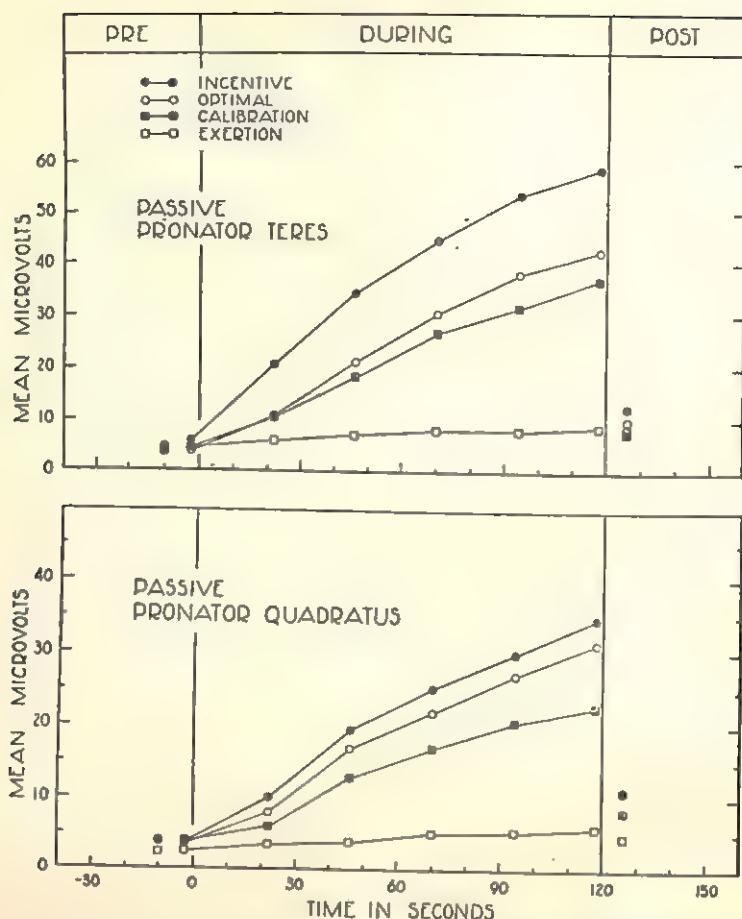


FIG. 2. Graphs from Stennett's experiment (37), showing mean EMG gradients obtained under conditions varying in degree of incentive. Steepness of gradient varies directly with degree of incentive.  $N = 31$ .

If we consider our physiological measures as indicants of arousal level, we may say that performance suffered in the first instance because of under-arousal (or poor motivation), while in the second instance it suffered from overarousal (or emotional interference). In short, as Stennett has previously stated (37), we believe that the concept of arousal leads us in the direction of working out (empirically) a continuum of behavioral intensity which promises to have the very desirable feature of integrating the concepts of motivation and emotion. From available data it appears that physiological measures, such

as palmar skin conductance, EEG<sup>4</sup> and gradients in skeletal muscle tension, heart rate, blood pressure and respiration (26) should provide reliable measures of the arousal variable. The objective nature of the physiological measures is a highly desirable feature which frees the worker from dependence upon merely manipulating situations in the

<sup>4</sup> Stennett (38) has found that the relationship of alpha amplitude to arousal level is nonlinear. On the lower end of the arousal continuum the relationship is positive, such that raising arousal leads to increasing alpha amplitude; but past the middle range of arousal the relationship becomes inverse. This latter function is the better known one.

hope that he is producing intended changes in the arousal level of his subjects. Moreover, the physiological indicants have the further advantage that they may be applied to work with animals as well as with human subjects, and may thus serve usefully to bridge the gap, in the field of motivation, between work on human and on infrahuman subjects.

A word should be said concerning the different physiological measures which have served as indicants of behavioral arousal. Although gradient steepness has proved a very useful measure, level of palmar skin conductance seems equally promising. As a matter of fact, even with EMG, the correlation between average EMG level and gradient steepness is usually so high that it is meaningless to ask which is a better indicant. We still have much to learn concerning the application of physiological techniques to our problems. It may be that, as Lacey's work suggests (23), for most accurate assessment of arousal, special consideration should be given to individual differences in relative reactivity of different physiological systems.

Following the usage of Freeman (15) and Hebb (18), the term "arousal" is used to refer to the intensive dimension. I am aware that the term "arousal" is used by some EEG workers to denote flattening of an EEG tracing (e.g., 8, p. 132). When I use the term, I use it in a much broader sense, as a dimension of behavior, and I am not using this term to refer to the EEG phenomenon called "arousal" or "activation." It is for this reason that I specify *behavioral* arousal in the title of this paper. As investigative work proceeds, it may become heuristic to make a definite distinction between physiological arousal and behavioral intensity. Granting this possibility, I believe that for present purposes it may be preferable to accept a rather broad operational definition of the in-

tensity dimension, in which level of physiological activity, arousal, and intensive level are employed as roughly synonymous terms.

In short, the physiological measures appear to be useful tools in establishing and precisely quantifying a dimension of behavioral intensity. Indeed, I regard such objective measures as nearly indispensable to the achievement of a really satisfactory operational definition of behavioral intensity. In the absence of such objective measures, it is difficult to see how circularity can be avoided. Considerable work is required, of course, in working out the intensity dimension, and while present results are indeed encouraging, many further data are required. It may be helpful just here to relate the arousal continuum to the intensity dimension which Boring described (4). While Boring's main concern was with sensation, I believe that it is appropriate to consider that operations of measurement comparable in precision to those of psychophysics may be possible in the field of action.

#### EXPERIMENTAL STUDIES WITH PSYCHIATRIC PATIENTS

Having elucidated the concept of arousal with these reference experiments, we are now in a favorable position to take a fresh look at the data comparing patients and nonpatients with respect to level of physiological reaction under controlled stimulating conditions. At the outset, we may say that the chief impression which one gets in going over all of these data is that, under "stress," psychoneurotic patients appeared to show a higher level of physiological reaction than controls, and that level of reaction seemed particularly high in patients suffering mainly from pathological anxiety. By pathological anxiety, I mean a state of such severity that work efficiency is seriously affected over long periods of time, and a state which is



characterized by one or more of the following complaints: persistent feelings of "tension" or "strain," "irritability," "unremitting worry," "restlessness," "inability to concentrate," "feelings of panic in everyday-life situations." I should like to make it very clear that I do not employ the term "anxiety" to refer to transient affective states. When I use the term I am talking about a pathological condition which—as far as we can determine—develops slowly, over months or years, and from which recovery (when it occurs) is also slow and gradual. The experiments which we shall consider in this section employed patients suffering from "pathological anxiety," as we have just defined it. For the sake of convenience, these subjects will be called "anxiety patients."

In a study with pain as standard stimulus (27), the following physiological measures showed significantly greater reaction in anxiety patients than in other psychiatric patients: finger movement (and number of voluntary pressures to indicate pain), neck-muscle activity, deviation in amplitude and rate of respiration throughout the test, respiratory irregularities occurring at time of stimulation, and heart-rate variability. In a different study (29) with a perceptual test and a Luria-type recording from the left hand, finger movement was significantly more irregular in anxiety patients than in other psychiatric patients.

To repeat an earlier statement, these findings indicate that under standard conditions of stimulation psychoneurotics are more reactive than controls, and that patients with anxiety predominating in the symptom picture are the most responsive of all.

#### *Need for "Standard Stress" in Demonstrating Differences Between Patients and Controls*

Another question which we sought to answer was whether a certain level of

arousal must be reached in order to demonstrate differences between patients and nonpatients or whether such differences could be obtained under resting, "basal" conditions. From reviewing the literature prior to conducting our own experiments, we were led to suspect that some stimulation would be necessary because experiments which had been carried out under resting conditions had usually yielded negative or inconclusive results.

Our findings did indeed clearly show that, in differentiating between patients and controls, some form of stimulation was definitively superior to merely taking records under resting conditions. This has been demonstrated for blood pressure (28, p. 89), for muscle potentials in motor tasks (31, p. 54 and pp. 59 ff.), and again for muscle potentials in two separate investigations of startle (30, p. 327; 7, p. 181). The only measure which we have found to discriminate well between patients and controls under "resting" conditions was frontalis-muscle potentials (33). However, we know that "resting" conditions associated with a testing session are by no means basal, and that—for example—significantly lower blood-pressure readings may be obtained from patients resting quietly on the ward than in the so-called "resting" condition of our experiments (32).

*"Specific" vs. "nonspecific" stimulating situations.* In producing higher levels of arousal in patients, is it necessary to present material to which patients are specifically sensitized or is it possible to demonstrate the difference between patients and controls by employing the same standard stimulating situation for all subjects? Our experiments clearly show that the latter is true. It is not necessary to present the patients with words or situations which have special meanings for them in order to produce more arousal in them than in controls.

As an example of a "specific-com-

plex" technique of producing high-level arousal, Luria (25) employed the method of controlled association in which he compared motor reaction to "critical" words (those which were especially arousing for the subject because of their association with specific life experiences) with reaction to indifferent words. Our situations, on the other hand, were chosen for their general arousal value, and we sought to avoid situations which would have special meaning for particular individuals.

With this point especially in mind we devised our standard situation of painful stimulation, because of the nearly universal avoidance reaction to pain. In order to permit more generalized conclusions, we also employed standard situations other than pain. One study is of especial interest because we reproduced the essential features of Luria's procedure, only substituting a series of size discriminations for the series of verbal stimulations which Luria employed (29). Conclusions from these experiments were as follows. All measures of motor activity recorded during performance of speeded size discrimination yielded reliable differences between patients and controls. In every instance there was evidence of greater physiological disturbance in the patients. The measures employed may be distinguished as skeletal-motor (motor control, muscular tension) and autonomic (systolic blood pressure). These differences in motor activity were manifested even though psychoneurotics, acute psychotics, and controls were practically identical with respect to perceptual performance.

These results led us to question certain views concerning determinants of higher arousal levels in psychoneurotics. In much current writing there is the underlying assumption that physiological disturbances in the psychoneurotic can be accounted for entirely in terms of situational explanations. These writers

assume that there is no need to look for pathology in central and motor mechanisms, because they believe that amount of physiological disturbance is commensurate with the special significance which the situation has for the patient. Implied in this view is the assumption that only those stimuli which, through learning, have acquired special meaning for the patient have the power to produce an "abnormal" level of arousal. It assumes that the patient may participate in many situations without showing abnormally high levels of physiological reaction.

However, this view may well be questioned because it does not appear to fit with clinical observations. Cameron has written as follows:

It will be noted that nearly all such patients [with anxiety states] complain that they cannot go into crowded places or into any situation where sustained efforts will be required of them. Their symptoms are made more severe by anything which elicits emotional reactions, such as altercations or participating in a discussion of illness. Nearly all find, at least at first, that their symptoms are increased by visiting their former places of employment or meeting fellow-workers. In other words, their symptoms are exacerbated by anything which serves to increase tension. *Emphasis should be placed upon the fact that their symptoms are elicited or intensified, not primarily by the reactivation of any conflict situation which may exist, but literally by everything in the course of the day which serves to increase tension* (5, pp. 56-57. Italics mine).

In therapy, relaxants of various kinds are devised to "damp" the "autonomous" reaction before proceeding with psychotherapy (41).

*Strong auditory stimulation.* Strong auditory stimulation served as another and very different kind of standard stimulating situation for comparing patients and controls. Two separate studies, the first one (30) with induced tension (produced by squeezing a rubber bulb), and the second (7) without induced tension and with a less intense stimulus, agreed in showing that the



most reliable difference between anxious patients and controls was in "after-response" following the period of primary reflex-startle reaction.

#### NEUROPHYSIOLOGICAL CONSIDERATIONS

In the interpretation of our findings in the experiments on strong auditory stimulation (7, 30), we cited the parallel between these observations on patients and findings in neurophysiological experiments on the reticular formation. In certain animal preparations, after-discharge in the cerebral cortex was abolished by stimulation in the reticular formation of thalamus (20) and brain stem (35). We believe that it is reasonable to suggest that some such inhibitory mechanism (as the one which abolished after-discharge) may be weakened in pathological anxiety.

Having implicated inhibition, we are required to examine this concept critically for a moment. Although there is by no means complete agreement on the matter of inhibitory mechanisms in the central nervous system, present evidence appears to point more and more in the direction of inhibition as a phenomenon in its own right, independent of excitation (i.e., not merely absence of excitation).

Of the current theories of inhibition known to me, Eccles' view seems most reasonable (12). Eccles and his co-workers developed a technique for placing a microelectrode within a single spinal motoneurone, and they were thus able to observe the electrical potential between the inside and the outside of the cell. They observed that when they stimulated an inhibitory nerve fiber it increased the polarization of the nerve cell on which it ended. Eccles called this effect "hyperpolarization," which, electrically is the opposite of what occurs when a nerve cell is fired (depolarized).

While Eccles' work was done on cells in the spinal cord, it nonetheless seems reasonable to suggest that the reticular formation could produce widespread inhibition in the cortex by hyperpolarizing cortical cells. Because the study of neuronal discharge in the cortex is a new field of research, sufficient data to decide this point are not at hand. But data which are presently available seem to be in line with the proposition that some impulse arriving in the cortex may have facilitatory effects, while others may produce opposite results (21, Fig. 19, p. 62).

If Eccles' theory is essentially correct,<sup>5</sup> we may work with inhibition as an independent process, and seek to understand the pathology of anxiety in terms of weakened inhibition. To make matters more concrete, we may draw on Eccles' hypothesis of a chemical transmitter for inhibition (12, p. 163) and on the recent experimental work of Elliott and Florey (13) to suggest that, in anxiety, the effectiveness of this substance has been reduced.

#### THE PROBLEM OF ETIOLOGY

The disorder of pathological anxiety may be conceived of almost entirely in terms of constitutional factors. It is logical to consider that certain individuals may inherit a deficient inhibitory mechanism. Such a person would consistently suffer from inability to relax throughout life, and would be seriously limited in the amount of stimula-

<sup>5</sup> Recent findings, although supporting Eccles' main conclusions, suggest that the phenomenon may be somewhat more complex than he originally supposed. The observations of Kuffler and Eyzaguirre (22) on inhibition of stretch receptor organs in crustaceans indicate that the polarity of the "inhibitory potential" varies with the state of the cell. When the cell is depolarized, an inhibitory volley causes polarization; when the cell is resting, an inhibitory volley causes depolarization.



tion that he could withstand. In such a case the constitutional weakness, rather than learning, would be the primary factor in etiology. While constitutional differences of genetic origin may account for degree of predisposition to the pathological condition of anxiety, clinical evidence stands against a purely genetic etiology. The fact that such a large number of patients recover from anxiety states (17, 34) argues against a purely genetic-constitutional explanation of pathological anxiety.

Declining the genetic-constitutional explanation of anxiety implies that learning mechanisms are somehow involved in the pathology. In order to understand the full implications of this point of view, it is helpful to consider that degree of arousal is not a "given" in the stimulating situation. The same stimulating situation may produce quite different levels of physiological reaction in different persons, depending upon the effects of past learning.<sup>6</sup> We may compare individuals with respect to their physiological reactions in a large number of different situations. We may find, for example, that a certain person generally shows significantly higher levels of physiological reaction than most other individuals. If this person can avoid stimulating situations with high-arousal values he appears no different from others. However, in ordinary, everyday living, it is unlikely that he will be

able to avoid such situations, and he will, therefore, be more or less constantly operating at physiological levels which are higher than normal. We may conjecture that in such a case in which stimulation keeps physiological levels constantly very high, over a long period of time there will be a weakening of inhibitory mechanisms from overuse.

#### FURTHER CLINICAL-EXPERIMENTAL CONSIDERATIONS

*Anxiety in combat.* If our theory is correct, anxiety may be considered as a "disease of overarousal" (or in Selye's [36] terms, a disease of "adaptation"). That is, the critical neural change is thought of as being produced by a process of attrition from excessive and extended overarousal. It would not matter whether this overarousal were produced in an individual whose previous learning made him more prone to overarousal, or whether the individual were anxiety-resistant from past training, and was simply "overexposed" to situations (like battle) that everyone reacts to with extremely high physiological levels. With this view we can readily understand why under battle conditions each soldier would have his "breaking point," and why despite resistance to overarousal from constitution and previous learning, if situations of high-arousal level are repeated over a long enough time period, the critical change will finally occur. This seems to be the picture which emerges from studies of anxiety in combat (16, pp. 85 ff.).

#### *Inhibitory Deficiency in Anxiety and in Manic States*

From the clinical point of view, Cameron (6, p. 388) has drawn attention to the prominence of overactivity in the anxiety states. Cameron is inclined to believe, however, that the manic state best represents "pure overfacilitation," in comparison with anx-

<sup>6</sup> The reader will recall that in our physiological studies of psychiatric patients we attempted to avoid experimental situations which had special meanings for particular individuals. In an earlier section of this paper we referred to these situations as "nonspecific." We assume that an anxiety-prone individual, before he actually develops the pathological state (and after he recovers from it), will not show higher arousal levels in such "nonspecific" stimulating situations. The stimulating situations referred to as producing quite different levels of physiological reaction in different persons are, of course, what we called "specific" in the earlier section of this paper.

iety, which he has described as "curbed overactivity." In drawing this comparison, Cameron was influenced by his careful observation of body movements. He found that the typical anxious patient was restless and in constant movement, but that he did not have the open, wide, flung-out movements of the manic. In general, the movements of the anxious patient remained within the body silhouette.<sup>7</sup>

The internally generated manic overactivity ("pure overfacilitation") could reasonably be accounted for by positing increased activity of facilitatory mechanisms.

#### PROBLEMS FOR FURTHER STUDY

The line of reasoning followed in the present paper suggests certain hypotheses which might be put to experimental test. In the first place, longitudinal physiological study of patients suffering severe states of anxiety should reveal changed physiological reaction under conditions of standard stimulation. That is, during performance of a motor task—for example, palmar skin conductance—electromyographic gradients and other physiological indicants of arousal should show decline when the patient is in remission, and should show increase again with relapse and return of the anxiety. This is a straightforward kind of investigation which one might suppose had already been under-

taken. However, as far as I am aware, the study has not been carried out with anxiety states in the way proposed.

#### *Anxiety and Learning*

Physiological measures of arousal should prove valuable in learning experiments in which anxiety has been studied as a variable (14). For example, workers have employed questionnaires and scales (e.g., the Taylor scale [40]), to select subjects high in "anxiety." The chief purpose of such experiments has been to compare the learning speed of subjects scoring high on such a scale with other subjects scoring lower on the scale. It would appear that physiological measures could be applied to such problems with considerable advantage. Subjects who would probably react at high physiological levels could still be selected with the scales as an initial screening device; but physiological measurements could then be applied to provide actual values to place each subject on a continuum. Such methodology would appear promising in providing a continuous variable (i.e., physiological intensity, or arousal) for study in place of the rather dubious anxiety-nonanxiety dichotomy, and would have other advantages. For example, a low scorer on the scale might be temporarily upset, and so be misclassified in an experiment unless his actual physiological measures were available on the day of the experiment.

#### *Research with Reserpine and Chlorpromazine*

<sup>7</sup> On the surface, this appears incongruous with the notion of weakened inhibition. However, we may account for this constrained appearance of inhibition by suggesting the substitution of less efficient mechanisms of inhibition for the one which has suffered impairment. It may be, for example, that anxiety patients compensate for weakened autonomous mechanisms by calling on voluntary motor mechanisms (i.e., the pyramidal motor system). For example, in the absence of sufficient control from autonomous inhibitory mechanisms, the anxiety patient may avoid loss of motor control through co-contraction of antagonistic muscles.

Patients exhibiting anxiety as the predominant symptom have been reported to improve significantly following the administration of reserpine and chlorpromazine (19). It should prove illuminating to study the effects of such drugs on physiological reaction of anxiety patients under controlled stimulating conditions. For example, with administra-



tion of these drugs, would the electromyographic reaction of patients to strong auditory stimulation resemble the normal reaction more closely (show less after-response) than in the absence of the drugs?

It would likewise be of interest to determine the effect of such drugs on levels of physiological reaction in anxiety patients under conditions of moderate stimulation, such as those in our experiments with pain and with performance tasks. Would drug administration bring levels of physiological reaction down close to normal values under these conditions?

As a matter of fact, our experiments with psychiatric patients were performed prior to the full development of the concept of an intensity continuum in behavior, measured in terms of EMG gradients, level of palmar skin conductance, and other such physiological indicators. It would be highly desirable, therefore, to apply these more refined physiological measures to the study of anxiety patients. Do they, in fact, show steeper EMG gradients than normals in tracking, and are these gradients reduced in slope with administration of reserpine and chlorpromazine?

#### *Proposed Animal Experiments*

Certain aspects of these problems may be more advantageously studied with animal subjects. Studies of "experimental neuroses," as reviewed by Liddell (24), have shown that it is possible to produce chronic states characterized by physiological deviation. For present purposes it would be desirable to employ a form of stimulation which effectively maintains high levels of physiological reaction over long periods of time. For our purposes it would not matter particularly how the stimulation was produced; the main requisite is that high physiological levels be recorded continuously over days and weeks.

The main purpose of such an experiment would be to determine whether keeping physiological levels constantly high would finally produce "anxiety" in animals (i.e., animals with raised physiological levels in standard test situations). If such experiments did turn out positively, valuable animal "preparations" would be available for neurophysiological and pharmacological studies.

Such a "preparation" might be used, for example, to determine whether inhibitory effects from stimulation in the reticular formation are weaker than in normal animals. We might even conceive of an experiment paralleling the ones which we carried out with human subjects. It would seem possible to implant electrodes in the reticular formation to search for areas which fire inhibitory impulses to the cerebral cortex following strong auditory stimulation. Furthermore, pharmacological investigation (13) might be directed to the question whether there is an inhibitory substance in the brain which becomes dilute with long-continued overarousal.

#### SUMMARY

The main purpose of this paper is to consider some recent experimental data which suggest a way out of the present confusion surrounding the concepts of motivation, emotion, and anxiety. Two lines of investigation, each employing physiological methods, are examined. In one experimental program, measures such as steepness of muscle-potential gradients and level of palmar skin conductance were found to be useful indicators of arousal level. The results of several experiments demonstrated significant relationships between such physiological indicants and excellence of performance on various motor tasks, such as mirror tracing and tracking. In this empirical setting, problems of relationship between concepts of motivation and emotion are reconsidered.



The arousal concept is then applied to the problem of pathological anxiety in psychiatric patients. The earlier results from physiological studies carried out with psychiatric patients as subjects are reviewed in the light of the more recent work on physiological indicants of arousal. Considerable confusion has arisen because the term "anxiety" has been used to denote two quite different states of the organism: (a) any increase in level of arousal, however brief the rise (or however selective the stimulating condition); and (b) a pathological state in which the patient appears chronically overreactive (physiologically) to every stimulating situation.

It seems reasonable to restrict the term "anxiety" to the chronic pathological condition. Results from physiological studies carried out with patients suffering this pathological condition indicated that standard stimulation (or "stress") accentuated the differences in arousal between anxiety patients and controls. Under resting conditions such differences were usually insignificant. On the basis of the data reviewed, certain hypotheses concerning the nature and etiology of pathological anxiety are tentatively advanced. It is suggested that anxiety may be produced in an individual (in animal as well as in man) by keeping level of arousal very high over long periods of time. Finally, recent neurophysiological findings are cited in stating the hypothesis that such continuous overarousal may result in impairment of central inhibitory mechanisms.

#### REFERENCES

1. BARTOSHUK, A. K. Electromyographic gradients in goal-directed activity. *Canad. J. Psychol.*, 1955, 9, 21-28.
2. BARTOSHUK, A. K. Electromyographic gradients as indicants of motivation. *Canad. J. Psychol.*, 1955, 9, 215-230.
3. BÉLANGER, D. J. "Gradients" musculaires et processus mentaux supérieurs. *Canad. J. Psychol.* (in press).
4. BORING, E. G. *The physical dimensions of consciousness*. New York: Century, 1933.
5. CAMERON, D. E. Autonomy in anxiety. *Psychiat. Quart.*, 1944, 18, 53-60.
6. CAMERON, D. E. Some relationships between excitement, depression, and anxiety. *Amer. J. Psychiat.*, 1945, 102, 385-394.
7. DAVIS, J. F., MALMO, R. B., & SHAGASS, C. Electromyographic reaction to strong auditory stimulation in psychiatric patients. *Canad. J. Psychol.*, 1954, 8, 177-186.
8. DELAFRESNAYE, J. F. (Ed.). *Brain mechanisms and consciousness*. Springfield, Ill.: Thomas, 1954. (See especially discussion by H. H. Jasper, p. 132.)
9. DUFFY, ELIZABETH. The conceptual categories of psychology: a suggestion for revision. *Psychol. Rev.*, 1941, 48, 177-203.
10. DUFFY, ELIZABETH. A systematic framework for the description of personality. *J. abnorm. soc. Psychol.*, 1949, 44, 175-190.
11. DUFFY, ELIZABETH. The concept of energy mobilization. *Psychol. Rev.*, 1951, 58, 30-40.
12. ECCLES, J. C. *The neurophysiological basis of mind*. Oxford: Clarendon, 1953.
13. ELLIOTT, K. A. C., & FLOREY, E. Factor I—Inhibitory factor from brain. Assay. Condition in brain. Stimulating and antagonizing substances. *J. Neurochem.*, 1956, 1, 181-192.
14. FARBER, I. E. Anxiety as a drive state. In M. R. Jones (Ed.), *Nebraska Symposium on Motivation*. Lincoln: University of Nebraska Press, 1954.
15. FREEMAN, G. L. *The energetics of human behavior*. Ithaca, N. Y.: Cornell University Press, 1948.
16. GRINKER, R. R., & SPIEGEL, J. P. *Men under stress*. Philadelphia: Blakiston, 1945.
17. HARRIS, A. The prognosis of anxiety states. *Brit. med. J.*, 1938, 2, 649-664.
18. HEBB, D. O. Drives and the C.N.S. (conceptual nervous system). *Psychol. Rev.*, 1955, 62, 243-254.
19. HOLLISTER, L. E., TRAUB, L., & BECKMAN, W. G. Psychiatric use of reserpine and chlorpromazine. Results of double-blind studies. In N. S. Kline (Ed.), *Psychopharmacology*. Washington, D. C.: Amer. Assoc. for Advancement of Science, 1956.
20. JASPER, H. H. Diffuse projection systems: the integrative action of the thalamic

- reticular system. *EEG Clin. Neurophysiol.*, 1949, 1, 405-420.
21. JUNG, R. Neuronal discharge. *EEG Clin. Neurophysiol.*, 1953, Suppl. No. 4, 57-71.
  22. KUFFLER, S. W., & EYZAGUIRRE, C. Synaptic inhibition in an isolated nerve cell. *J. gen. Physiol.*, 1955, 39, 155-184.
  23. LACEY, J. I. Individual differences in somatic response patterns. *J. comp. physiol. Psychol.*, 1950, 43, 338-350.
  24. LIDDELL, H. S. Conditioned reflex method and experimental neurosis. In J. McV. Hunt (Ed.), *Personality and the behavior disorders*. New York: Ronald, 1944. Vol. I, pp. 389-412.
  25. LURIA, A. R. *The nature of human conflict*. New York: Liveright, 1932.
  26. MALMO, R. B., & DAVIS, J. F. Physiological gradients as indicants of "arousal" in mirror tracing. *Canad. J. Psychol.*, 1956, 10, 231-238.
  27. MALMO, R. B., & SHAGASS, C. Physiologic studies of reaction to stress in anxiety and early schizophrenia. *Psychosom. Med.*, 1949, 11, 9-24.
  28. MALMO, R. B., & SHAGASS, C. Studies of blood pressure in psychiatric patients under stress. *Psychosom. Med.*, 1952, 14, 82-93.
  29. MALMO, R. B., SHAGASS, C., BÉLANGER, D. J., & SMITH, A. A. Motor control in psychiatric patients under experimental stress. *J. abnorm. soc. Psychol.*, 1951, 46, 539-547.
  30. MALMO, R. B., SHAGASS, C., & DAVIS, J. F. A method for the investigation of somatic response mechanisms in psychoneurosis. *Science*, 1950, 112, 325-328.
  31. MALMO, R. B., SHAGASS, C., & DAVIS, J. F. Electromyographic studies of muscular tension in psychiatric patients under stress. *J. clin. exp. Psychopath.*, 1951, 12, 45-66.
  32. MALMO, R. B., SHAGASS, C., & HESLAM, R. M. Blood pressure response to repeated brief stress in psychoneurosis: a study of adaptation. *Canad. J. Psychol.*, 1951, 5, 167-179.
  33. MALMO, R. B., & SMITH, A. A. Forehead tension and motor irregularities in psychoneurotic patients under stress. *J. Personality*, 1955, 23, 391-406.
  34. MILES, H. H. W., BARRABEE, EDNA L., & FINESINGER, J. E. Evaluation of psychotherapy. *Psychosom. Med.*, 1951, 13, 83-105.
  35. MORUZZI, G., & MAGOUN, H. W. Brain stem reticular formation and activation of the EEG. *EEG Clin. Neurophysiol.*, 1949, 1, 455-473.
  36. SELYE, H. *Stress*. Montreal: Acta, 1950.
  37. STENNETT, R. G. The arousal continuum. *J. exp. Psychol.* (in press).
  38. STENNETT, R. G. The relationship of alpha amplitude to the level of palmar conductance. *EEG Clin. Neurophysiol.*, 1957, 9, 131-138.
  39. SURWILLO, W. W. Psychological factors in muscle-action potentials: EMG gradients. *J. exp. Psychol.*, 1956, 52, 263-272.
  40. TAYLOR, JANET A. A personality scale of manifest anxiety. *J. abnorm. soc. Psychol.*, 1953, 48, 285-290.
  41. TYHURST, J. S., & RICHMAN, A. Clinical experience with psychiatric patients on reserpine—preliminary impressions. *Canad. med. Assoc. J.*, 1955, 72, 458-459.
  42. WALLERSTEIN, H. An electromyographic study of attentive listening. *Canad. J. Psychol.*, 1954, 8, 228-238.

(Received December 10, 1956)

## OPTICAL MOTIONS AND TRANSFORMATIONS AS STIMULI FOR VISUAL PERCEPTION<sup>1</sup>

JAMES J. GIBSON

*Cornell University*

How do we see the motions of objects in the world around us? The way to go about answering this question is to note the kinds of physical motion that occur in the human environment and then to examine the kinds and variables of optical stimulation that correspondingly occur. The isolation and control of these variables with suitable optical apparatus will make possible an experimental psychophysics of kinetic impressions. The desirability of such a psychophysical approach has been pointed out in an earlier paper (4) and the following proposals modify or make explicit a number of suggestions there made.

### DISTINCTION BETWEEN PHYSICAL MOTIONS AND OPTICAL MOTIONS

Ever since Isaac Newton supposed that the motions of things revealed the forces behind them and thereby the causes of all events, physics has been concerned with the observation of motions. The beauty of the idea for physics is that it applies to *all* things: stars and planets; stones, machines, and animals; particles and atoms. Of these motions, however, only a certain class is the concern of perceptual psychology. The things whose motion is visible are substances which, in the first place, dif-

ferentially reflect or emit light and, in the second place, are not either too far away or too small. The motion of the wind is invisible because gaseous substances transmit light instead of reflecting it. The motion of the heavenly bodies is invisible because their angular change of position is too small per unit of time. And the motion of microscopic bodies is invisible because their boundaries reflect an optical texture too small for the eye to resolve. But the environment of man contains an enormous variety of surfaces which do project focusable light to his eye, and these are the bodies the psychologist must be interested in. For when they fall, rise, turn, roll, bend, flow, twist, writhe, stretch, or break, an eye can register this event and the animal possessing the eye can respond to it.

We may observe that physical motions can be classified as *rigid* or *non-rigid*, the former being characteristic of crystalline substances and the latter of elastic or fluid substances. Rigid physical motions are exemplified in mechanics; they are analyzable into components of translation and rotation on or around any of three axes, and they have been studied since the time of Newton. Non-rigid physical motions are exemplified in biology since the growth and also the reactive movements of living animals are generally of this sort (10).

The motions of the physical environment might also be classified as *connected* or *nonconnected*. In the former the parts of the moving substance remain adjacent, even if not rigid, whereas in the latter they do not and are not even considered parts of the "same"

<sup>1</sup> This essay is part of an address entitled *Stimulation and Perception* delivered by the author in September 1955 at San Francisco as retiring president of the Division of Experimental Psychology, APA. The film which accompanied it, and serves to illustrate this paper (5), was produced with the support of the Office of Naval Research under Contract NONR 401(14) with Cornell University, which is concerned with research on the perception of motion and space.



substance. Instead they are treated as separate motions of different objects. The separations and fusions, attractions and repulsions, or collisions of things, animals, and people are all of this sort.

Physical motions are given to an eye only in the form of optical motions. An optical motion is an event in the *optic array*, that is, in the light reflected from an illuminated environment to an eye or, rather, to any position in the air where an eye might be placed (7). An eye is an organ for exploring an optic array. The solid sector of this array which an eye takes in is the basis of patterned vision; neither objects nor their motions could affect the eye except by means of it. External motion can be seen only if some differentiated part of the array is displaced relative to the rest of it, or to some other part, or if parts move relative to one another. There has to be some change of its pattern, considered as a projection to a point.

An optical motion, then, is a projection in two dimensions of a physical motion in three dimensions. The projection may be taken either as on the surface of a sphere centered at the eye or as on a plane in front of the eye. When locomotor movements of the observer are to be considered, the former is preferable (8), but when, as here, they are not involved, the plane projection is better. The one can be converted into the other if necessary. Our question is, What kinds of optical motion occur which might serve as stimuli for perception?

#### KINDS AND VARIABLES OF OPTICAL MOTION

How can optical motions be described or specified? The question has to do with the motions of a texture or pattern in a two-dimensional array. Tentatively, there seem to be two general

possibilities. First, one could divide the pattern into convenient elements, describe the positions of all the elements by pairs of coordinate values, and finally describe the motions of all the elements by the successive pairs of values. Or one could describe the motions of the elements by direction and speed at successive moments of time. This procedure is analytical. Second, a non-analytical method of specifying optical motions is possible. One could simply take the pattern as given, and then use the operation defining a *perspective transformation* in geometry to describe a family of changes of pattern. This method does not exhaust all the varieties of optical motion, as will be evident, but it has advantages for an experimenter who needs to produce an artificial optical array.

*Continuous perspective transformations.* In geometry, any form or pattern on one plane which is a projection of a form or pattern lying on some other plane is called a *perspective transformation*. These forms are static. When the point of projection (the focus of the sheaf of lines which connect the pair of forms, point for point) is near the two planes, we speak of a *central* or *polar* projection; when this point of projection is at a very great distance from the planes, we speak of a *parallel* projection. (It may be useful to recall that the "plans" and "elevations" of an architect are cases of parallel projection, but that his drawing for the client's eyes is a case of polar projection. This latter is the case we are chiefly concerned with.) When the two planes are parallel, the difference between the projected form and the given form is one of size only; it is called a *similarity* transformation. When the two planes are not parallel, the difference between the forms is that to which common meaning applies the term "perspective," or sometimes the term "foreshortening." In geometry,

both the difference of size and that of shape are classed as perspective transformations.

We can now speak of motion. When the angle or the distance of the first plane relative to the second plane is altered, the projected form is correspondingly altered. The fact is that the relation between any earlier and any later projected form is *also* a perspective transformation. The relation holds between any two of its stages in times. Hence, the motion in question may be described as a continuous series of perspective transformations. It is a relation between a temporal series or family of static forms, not merely between two forms. Any such moving transformation can be analyzed by six parameters corresponding to the six components of the possible movements of the first plane—that is, three of translation and three of rotation.

*Families of continuous perspective transformations.* It can now be observed that all optical motions resulting from the rigid physical motions of the flat face of an external object are continuous families of perspective transformations, as defined above. These are optical motions as taken on a plane in front of the eye. The six parameters of optical motion can be visualized as (a) vertical translation of the pattern in the plane, (b) horizontal translation of the pattern, (c) enlargement or reduction of the pattern, (d) horizontal foreshortening of the pattern, (e) vertical foreshortening of the pattern, and (f) rotation of the pattern in the plane.

These parameters of transformation are for the case of a *polar* projection. As the focus of projection is taken at an increasing distance from the two planes, one approaches the case of a *parallel* projection. For the latter case, three of the six parameters have been altered in character, namely (c), (d), and (e)

above, while (a), (b), and (f) remain unaltered. Enlargement or reduction of the pattern has vanished; horizontal foreshortening becomes a mere horizontal flattening; and vertical foreshortening becomes vertical flattening.

These "pure types" of optical motion can be observed on a motion picture screen (5). An irregular or regular contour shape or an irregular or regular group of spots can be made to undergo continuous perspective transformations, and the observer can note the various types of motion in the plane of the screen. One can note, for example, that in types (d) and (e) a square is transformed into a trapezium with polar projection, but is transformed into a flattened rectangle with parallel projection. The interesting fact, however, is that for types (c), (d), and (e) with polar projection it is *very difficult to notice* the motion in the plane. Instead, one sees a sort of "virtual" object or surface which (c) moves toward or away from the screen, (d) rotates on its vertical axis, or (e) rotates on its horizontal axis. One sees, in other words, rigid motion in depth. The suggestion is that the parameters of transformation are stimuli for the phenomenal parameters of the motion in space of one face of an object.

The rotations in depth are similar in some respects to the kinetic depth effect obtained by Wallach (14). Such effects have been observed for Lissajous figures (3), and long ago for the silhouette of a rotating windmill against the horizon (1, p. 270). All these apparent rotations are said to be characterized by ambiguity as to the direction of rotation, and by spontaneous reversals in the direction of rotation. The apparent rotations shown in the film, however, are *not* characterized by ambiguity or reversals of direction when the transformations are those obtained with polar

projection, but only when the transformations are those obtained with parallel projection.

A psychophysical experiment has been performed on the degree of perceived semirotation in depth as a function of the transformation sequence (6).

#### APPARATUS FOR PRODUCING CONTINUOUS PERSPECTIVE TRANSFORMATIONS IN THE OPTIC ARRAY

The method used to display the geometrical transformations on a motion picture screen was not by "animation" of film; the procedure, rather, was to photograph the window of a device which might be called a shadow transformer. Details of its construction are given in the report of the experiment (6). It will here be described only as it suggests possibilities for a psychophysics of motion perception. It consists of a translucent screen with a point source of light on one side and an observer symmetrically on the other side. In the square luminous window, dark shapes, patterns, or textures can be made to appear. They are shadows, not objects, so that only two grades of intensity exist, surface texture and binocular disparity are eliminated, and accommoda-

tion and convergence are controlled for this sector of the optic array to the eye. The variables of form and transformation are thus isolated for study.

In the diverging ray sheaf from the point source to the translucent window a mount can be placed, a pane of transparent material large enough so that it can be moved without its edges being projected on the screen. Forms, patterns, or textures cut from gummed paper or masking tape can be attached to this mount, or it can be traced with ink or even sprinkled with talcum powder, so that shadows of many varieties are projected on the screen. The mount can be rotated on any of three axes, or translated in any of three dimensions. Hence, considering the mount and the screen as two geometrical planes, changes in the position of the mount will yield all possible perspective transformations of the shadow relative to the shadow caster, and likewise all parameters of continuous perspective transformation of the shadow itself. Previous shadow-casting devices, most recently Wallach's (14), have not been constructed for this systematic purpose. They have also not utilized polar projection. For purposes of comparison, the present apparatus can also be illuminated by a projector

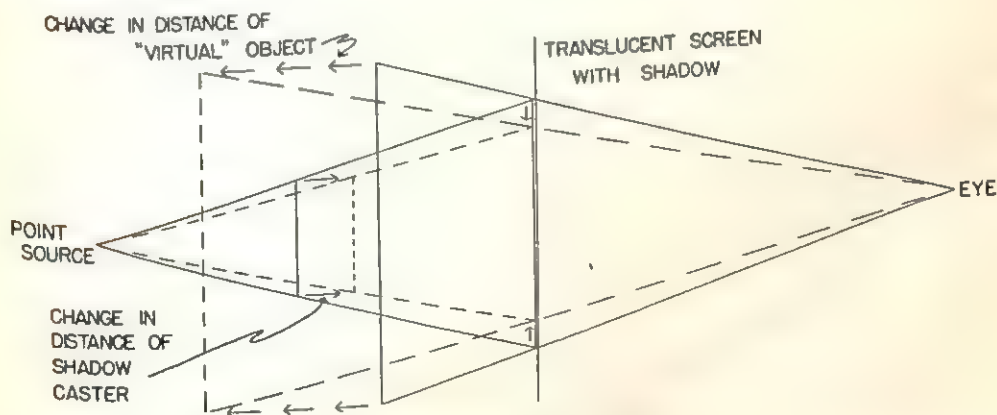


FIG. 1. The shadow transformer.



beam at 80 ft. instead of a point source at 5 ft.; the former yields an approximately parallel projection. The use of a transparent mount obviates the necessity of any visible support for the shadow caster, or of having the shadow it casts extend below the bottom of the window.

The optical geometry of the apparatus is given in Fig. 1 for a size transformation. It may be noted that the converging ray sheaf to the eye is the geometrical opposite of the diverging ray sheaf. The relation of the shadow to its "virtual object" (by analogy with the "virtual image" in a mirror) is simply the reverse of the relation of the shadow to the shadow caster. This reversal does not affect the transformations in any other way. It is true that the motions perceived are opposite to those of the shadow caster, but this does not present a paradox if one remembers that vision depends on the optic array, not on the way an optic array can be artificially produced.

#### INTERNAL DEPTH OF TRANSPARENT OBJECTS AND COLLECTIONS OF OBJECTS

So far, consideration has been limited to continuous transformations of textures corresponding to the opaque face of an object. What about the optical motions corresponding to the physical motions of transparent objects, or things with different parts in depth, or rigid collections of things like a forest of trees or a tangle of wire? The phenomenon termed "stereokinesis" (11, Ch. 13 and references) and the kinetic depth effect of Wallach (14) involve an impression of *internal depth*. The "virtual object" of these experiments is often a collection of posts or a figure of bent wire.

The complex optical motions of these experiments can be analyzed by considering different parameter values of the

same perspective transformation corresponding to the different planes of depth of the object. This suggestion becomes clearer with a concrete example. If, with the shadow transformer, one casts a shadow on the screen through several parallel sheets of glass, each of which has been sprinkled with talcum powder, the composite shadow looks something like the Milky Way in the night sky. When the sandwich of mounts is *moved*, however, the perception separates into clear planes of depth, each layer of nebulous material standing in front of its neighbor. The effect is shown in the film (5).

The internal depth of the virtual object produced by the moving shadow of a bent wire, or by an arrangement of vertical sticks on a horizontal turntable, has been frequently studied. The writer suspects that it will yield to the kind of analysis proposed. The depth is similar to binocular stereoscopic depth; a unifying hypothesis would be that the simultaneous disparity of binocular images is only a special geometrical case of the successive disparities of a continuous image. Both rest on the geometry of parallax, that is, the projection of a collection of objects in space to a point in space, and this has led Tschermak to include both under the term "parallactoscopy" (13). Both are transformations in the most general sense of the term, and perhaps both sorts of disparity should be treated as transformations.

*The impression of surfaces meeting at a corner.* With the apparatus described, one can also produce the impression of *more* than one flat face of an object moving in depth. If two transparent mounts are joined to make an angle, and if both are given some opaque texture of any kind, the combined shadow looks like a plane surface so long as the combined mounts are kept stationary. When they are moved, however, the shadow

becomes two faces or surfaces making a corner. The relative slant of one surface to the other can be judged with some accuracy.

#### WHAT IS THE STIMULATING EFFECT OF NONPERSPECTIVE TRANSFORMATIONS IN THE OPTIC ARRAY?

If the rigid mechanical motions of the physical environment are represented by one kind of geometrical transformations in the pattern of light, are the nonrigid biological motions of the environment represented by a *different* kind of geometrical transformations in the pattern of light? The difference is suggested when the geometer describes topology as "rubber sheet geometry." This is concerned with changes of bidimensional form *other* than the changes heretofore described. Considering an organism in silhouette, its reactive movements cannot be compounded of the six pure types of optical motion considered above. Neither can its growth be described as magnification. Medawar, a biologist following D'Arcy Thompson, seems to have demonstrated that the change of shape of the human figure from infancy to adulthood, disregarding the change of size, is a specific continuous transformation which can only be suggested in words. A "tapered stretch" describes it approximately (10, pp. 177 ff.). The change is monotonic, i.e., it keeps the same trend. Geometrically, there are different *forms* of change of bidimensional form. Conceivably, the visual mechanism is sensitive to these forms of change.

The shadow transformer can be adapted to display nonperspective transformations if an elastic or flexible sheet is used for the transparent mount which carries the shadow-casting form or texture, and if this is stretched or bent in some way. Preliminary observations suggest that the resulting perception of

motion is correspondingly elastic instead of rigid. There are technical difficulties in controlling such optical motions. But if apparatus can be built for systematically producing them, it will be open to the perceptual psychologist to study such phenomena as animate, expressive, and physiognomic movements in the manner of psychophysics.

#### DISJUNCTIVE OR SEPARATE OPTICAL MOTIONS

The converting of a single form on a plane into *two* forms is something which goes beyond the continuous transformations heretofore considered. There is, instead, a discontinuity in both the temporal and the spatial series. The geometer is tempted to describe it by saying that there is a breaking or tearing of the surface, thus falling back on a physical analogy.

If certain parts of a connected optical pattern undergo one kind of transformation while other parts undergo another kind, it might be predicted that the perceived surface will become two perceived surfaces, each composed of the parts carrying the same transformation. This is obvious when one set of parts moves in one direction and the other in a different direction, and the fact was recognized in Wertheimer's law of "common fate" as a determiner of sensory organization (11, Ch. 13). It should equally be true, however, when one set of parts carries a slant transformation or a size transformation different from that of the other; the texture will break into two textures each moving in its own tridimensional way. Some of these possibilities have been investigated by using a "sandwich" of mounts in the apparatus, and the film illustrates a few of these possible dual disjunctive motions. Perceptual separation does result. Evidently when the parts of an optical texture undergo *joint motion* this does not

have to be understood as a set of motions with the same velocity in the same direction.

It is also possible to note what happens when *all* the parts of a connected optical pattern move, each in a different direction: the pattern becomes many smaller objects, like a swarm of ants. This result also suggests that what connected the elements of the pattern in the first place was their nonmotion relative to one another; in the optic array, after all, stability is only a special case of transformation. Research on the problem of how elementary motions might be *organized* into a single unitary motion has been performed by Johansson (9), Duncker (2), and Metzger (11).

The varieties and dimensions of optical motion in which the parts are *not* connected in adjacent order are of formidable complexity. It is not even clear how to go about classifying them. Disjunctive optical motions are, however, the stimuli by virtue of which we see occurrences, happenings, and actions in the world around us. There is certainly order and lawfulness in them, for Michotte has studied the impressions of causality induced by higher order variables of nothing more objective than the motions of separate spots (12). These abstract variables are clearly discriminable by observers, for the impressions can be made to come or go as the experimenter varies certain spatiotemporal conditions.

### CONCLUSIONS

If the optical geometry here expounded is correct, there is a possible basis in optical stimulation for the ability to distinguish between and among rigid, elastic, and multiple moving things. The basis lies in different mathematical modes of transformation and motion. The implication is that we see both the constant and the changing properties of

things in this way. We see them not because we have formed associations between the optical elements, not even because the brain has organized the optical elements, but because the retinal mosaic is sensitive to transformations as such. These are stimuli for perception.

Is it really plausible, one might ask, to call anything as apparently abstruse as a continuous series of transformations a *stimulus*? A bit of evidence may here be convincing. A puff of air to the cornea of the eye is a stimulus for the blink reflex in the pure and original sense of the term. The fact is that when an observer with the apparatus described is near the screen, a rapid expansion of the shadow until it fills the screen will also produce a blink reflex. The latter event ought to be considered as much a stimulus as the former.

### REFERENCES

1. BORING, E. G. *Sensation and perception in the history of experimental psychology*. New York: D. Appleton-Century, 1942.
2. DUNCKER, K. Über induzierte Bewegung. *Psychol. Forsch.*, 1929, 12, 180-259.
3. FISICHELLI, V. R. Effect of rotational axis and dimensional variations on the reversals of apparent movement in Lissajous figures. *Amer. J. Psychol.*, 1946, 59, 669-675.
4. GIBSON, J. J. The visual perception of objective motion and subjective movement. *Psychol. Rev.*, 1954, 61, 304-314.
5. GIBSON, J. J. *Optical motions and transformations as stimuli for visual perception*. Motion picture film. Psychological Cinema Register, State College, Pa., 1955.
6. GIBSON, J. J., & GIBSON, ELEANOR J. Continuous perspective transformations and the perception of rigid motion. *J. exp. Psychol.*, 1957, 54, 129-138.
7. GIBSON, J. J., PURDY, J., & LAWRENCE, L. A method of controlling stimulation for the study of space perception: the optical tunnel. *J. exp. Psychol.*, 1955, 50, 1-14.
8. GIBSON, J. J., OLUM, P., & ROSENBLATT, F. Parallax and perspective during aircraft



- landings. *Amer. J. Psychol.*, 1955, 68, 372-385.
9. JOHANSSON, G. *Configurations in event perception*. Uppsala: Almqvist and Wiksell, 1950.
10. LE GROS CLARK, W. E., & MEDAWAR, P. B. *Essays on growth and form presented to D'Arcy Thompson*. New York: Oxford Univer. Press, 1945.
11. METZGER, W. *Gesetze des Sehens*. Frankfurt: Waldemar Kramer, 1953.
12. MICHOTTE, A. *La perception de la causalité*. Louvain: Publications Universitaires, 1954.
13. TSCHERMAK-SEYSENEGG, A. Über Paralaktoskopie. *Arch. f. d. ges. Physiol.*, 1939, 241, 454-469.
14. WALLACH, H., & O'CONNELL, D. N. The kinetic depth effect. *J. exp. Psychol.*, 1953, 45, 205-217.

(Received July 18, 1956)

## SOME PROBLEMS OF ECLECTICISM

MARY HENLE

*Graduate Faculty of Political and Social Science,  
New School for Social Research*

Some ten years ago Woodworth, commenting on the situation in psychology as a whole, wrote: "Some may lean toward one school and some toward another, but on the whole the psychologists of the present time are proceeding on their way in the middle of the road . . ." (21, p. 254). He suggested that "If we could assemble all these psychologists [all the psychologists in the world] in a convention hall and ask the members of each school to stand and show themselves, a very large proportion of the entire group would remain seated" (21, pp. 254-255).

A similar position was taken by Boring at about the same time:

During the 1930's the *isms* pretty well dropped out of psychology. . . . The only reason for mentioning these four schools in this book is that the student hears about *behaviorism* and *Gestalt psychology* and has a right to be told what they are and that they are no longer important as schools. What was good in all the schools is now simply part of psychology (4, p. 11).

The eclectics, rising above the conflict of schools, hold that psychologists today are in happy agreement. It is their position that no real issues exist among the various points of view in psychology. For it is only on the basis of such a belief that one is able to select from each approach, combining the theoretical contributions of all. As Woodworth presents the middle-of-the-road position:

Every school is good, though no one is good enough. . . . One points to one alluring prospect, another to another. . . . Their negative pronouncements we can discount while we accept their positive contributions to psychology as a whole (21, p. 255).

If eclecticism is as prevalent in contemporary psychology as the above statements suggest, it becomes important to examine its consequences for theory. It is the hypothesis of the present paper that the eclectics have, to a large extent, succeeded in reconciling differences only by obscuring theoretical issues. An alternative to this kind of eclecticism will be proposed.

We may begin with an examination of specific instances of eclectic reconciliations of differences. Examples of two kinds of eclecticism will be discussed, one having to do with reconciliations of positions which refer to the entire field of psychology—the attempt to resolve conflicts among "schools" or general points of view—and a more circumscribed eclecticism relating to particular psychological problems.

As an instance of the first kind, Woodworth<sup>1</sup> writes:

A broadly defined functional psychology starts with the question "What man *does*" and proceeds to the questions "How?" and "Why?" . . . So broadly defined . . . functional psychology scarcely deserves the name of a school because it would include so many psychologists who have not professed themselves. Now the question is whether our middle-of-the-roads are not after all members of this broadly conceived functional school. . . . But if the middle-of-the-roads are really functionalists, the question is then whether the same would not be true of all the schools. Are they not all functionalists at heart? (21, p. 255).

Commenting on such a functionalism, Boring wrote in 1950: "Woodworth be-

<sup>1</sup> A number of the examples to be considered will be taken from Woodworth's writings because the present author regards him as one of the clearest of the eclectics and one of those whose theories are to be taken most seriously.

lieved that psychologists were more in agreement than their quarrels indicated, and he sought a system to which all could subscribe. He very nearly succeeded (3, p. 565).

It is not likely to be denied that psychology today has a functionalist flavor. The interest in the adaptive value of psychological processes is everywhere apparent. Nor can there be much disagreement with a functionalism defined in terms of these three questions of Woodworth's. But it tells us very little about a psychologist to say that he is a functionalist in this sense. What we need to know are the kinds of answers which a particular psychology gives to these questions. These are the issues in contemporary psychology, and here it is that disagreements arise. What, for example, is the relation of reinforcement, or of repetition, to learning? How does the learning process proceed? What is the nature of the fundamental human motives? How does the group exert its influence on the individual? It is the answers to such questions as these that divide psychologists. It would seem that Woodworth has succeeded in bringing all together only by obscuring such issues. If all are functionalists today, we still have the problem of examining the differences among the several varieties of functionalism.

The point may be made more specifically in connection with the same author's theory of the conditions of transfer of training, the doctrine of identical components (19), reformulated by Woodworth and Schlosberg (23) as a theory of "common factors." It is put forward to resolve the differences between those who hold that transfer is a function of the identical elements in two learning tasks, and those who maintain that it is a matter of the application of common principles or other whole properties to the two activities. Wood-

worth's view<sup>2</sup> is that anything concrete can be transferred—thus both identical elements and principles—since "any idea that can be recalled, or any attitude that can be reinstated is concrete enough to qualify. Perhaps anything that can be learned can be transferred" (19, p. 207). Again, "what is successfully transferred is usually something you can put your finger on—a principle, a good emotional attitude, a technique" (22, p. 582).

The controversy about the conditions of transfer is settled, in other words, by saying that *something* is carried over from one activity to the other. It is true that this is a formulation which covers most of the cases. But it lacks an advantage of both of the theories it displaces, namely the attempt to state the specific conditions of transfer. It would seem that the differences are resolved only at the expense of any specific theory in the area of the controversy. The theory covers all the cases only by telling us nothing specific about any of them.

To return to the more general kind of eclecticism, there exist today a number of efforts to reconcile the various significant theoretical positions in psychology. A number of authors have attempted to resolve the differences between behavior theory and psychoanalysis, Gestalt psychology and psychoanalysis, behavior theory and Gestalt theory.<sup>3</sup> Several examples of this trend in contemporary psychology will be examined.

<sup>2</sup> Since the earlier formulation (19) is the more explicit, and since the later (23) seems not to differ from it in principle, the former will be drawn upon here.

<sup>3</sup> Not all of these discussions are eclectic, at least in the meaning used here. Some attempt to understand the contributions of one psychology in terms of the theory of another, rather than to resolve differences between them. Such work, for example *Personality and Psychotherapy* by John Dollard and Neal E. Miller, will not be considered here. It presents interesting problems of its own which deserve separate treatment.



The problems seen above in Woodworth's formulations exist also in Abt's statements about basic agreements between psychoanalysis and Gestalt psychology with respect to the structure and development of the personality:

Freud's multiple-structured self is not essentially different conceptually from Lewin's division of the person into regions. The dynamic and economic interchanges that are postulated as occurring with respect to the id, ego and superego in psychoanalysis find parallel expression in Lewin's system of barriers and the classes of movements across them (1, pp. 38-39).

If Abt means that Lewin's *metatheory* is compatible with Freud's *theory* of personality structure, a case could be (but has not been) made for this position. But to equate Lewin's division of the personality into regions with Freud's topographical analysis of the person is to lose all the specific psychological insights of the latter and much of the metatheoretical contribution of the former. Lewin's inner-personal regions, if translated into Freudian terms, would undoubtedly fail to distinguish between id, ego, and superego; his motor-perceptual region includes some but not all of the functions of the Freudian ego.<sup>4</sup> Lewin is, indeed, largely unconcerned with the specific content of the personality, with the distinction between conscious and unconscious motivation, and with the historical development of the person; it is impossible to discuss Freud's topographical divisions apart from such considerations. Again, Lewin has not discussed the content and nature of the forces responsible for behavior; thus the parallel with Freud's

statements about dynamics can be maintained only if one turns one's attention away from the specifics of Freud's successive instinct theories. It is indeed possible to find parallels between Lewin's statements about the tendency of systems under tension to seek discharge and Freud's formulations about the pleasure principle; but to equate the two is to lose the specific character of both the pleasure and reality principles and to neglect Freud's theories about behavior which is independent of the pleasure principle (8).

A further illustration will be given of the tendency prevailing in contemporary psychology to reconcile the ideas of Freud and Lewin. Another author writes: "It is the thesis of this paper that a synthesis of the ideas of Lewin and Freud provides a basis for the beginnings of an integrated system of psychological theory . . ." (5, p. 206). More specifically, "We have already noted the structural parallelism between Freud's divisions of the personality and Lewin's psychical systems" (5, p. 222). This writer, it is true, limits the parallelism, remarking, "Freud has provided the living clay for the Lewinian scaffolding" (5, p. 228).<sup>5</sup> Still, closer examination suggests that essential differences between the "living clay" and the "scaffolding" have been neglected. For example, "One of these [defense mechanisms], projection, becomes the equivalent for Lewin's unreality" (5, p. 222). While there are, of course, important components of unreality in projections, the two concepts are by no means equivalent, if only because not all events on a level of unreality (for example,

<sup>4</sup> For example: "It is to a certain degree arbitrary where one draws the boundary between the motor-perceptual system and the inner regions, whether for instance one considers the understanding of speech as an event within the boundary zone or within the inner-personal systems" (16, p. 178).

<sup>5</sup> In this connection Bronfenbrenner comments on Lewin's neglect of the content of psychical systems: "This is indeed an unfortunate oversight" (5, p. 214). The thesis will be developed elsewhere that this is no oversight, but that Lewin undertook a different task.

dreams, fantasies, vague hopes and wishes) can be described as projections, at least as Freud uses the term. Again, to put the Freudian unconscious "directly into Lewinian language" by saying that "there are sub-systems within the region of the self which are not in communication with each other" (5, p. 225) is to slip over the specific nature of unconscious processes in Freudian theory. Furthermore, if this is meant as an equation, it fails to do justice to the very important communications which do exist between conscious and unconscious systems. The point may be illustrated by the dream, which draws upon the person's waking experiences (e.g., experiences of the "dream day") and which is recalled by the waking individual.

Another instance of a premature reconciliation of Gestalt psychology and psychoanalysis may be taken from the work of Witkin *et al.* (18).<sup>6</sup> These authors point out:

Although psychoanalytic theory, in its conception of primary and secondary processes, recognizes the relation between intellectual functioning and personality, it has not really been concerned with the nature of secondary processes. . . . Gestalt psychology, in contrast, has offered a well-developed theory of cognition, in which the role of the nature of reality in determining perceptual and thought processes has been emphasized. But . . . Gestalt theory has on the other hand neglected the role of personal factors in perception. By showing that a perceptual act cannot be understood without reference to *both* personal factors *and* the nature of reality, studies such as ours help to bridge the gap between Gestalt and psychoanalytic theory, and provide a basis for bringing together the main aspects of both into a single comprehensive theory of human psychological functioning (18, p. 481).

If, as I believe, the authors are correct in saying that "there is still lacking

<sup>6</sup> These remarks in no way detract from the excellence or the significance of these authors' study.

in psychoanalytic theory any specific account of cognition or of the nature of secondary process" (18, p. 481), and that Gestalt psychology has neglected personal factors in perception—these are precisely the reasons why studies such as theirs do *not* help to bridge the gap between the two theories. The gap can be bridged only by a true reconciliation of existing differences; and since these two approaches have been concerned, as the authors point out, with such different areas of psychology, it is difficult even to know where the essential differences lie. A systematic analysis of the assumptions of both psychologies, one concerned with implicit as well as explicit assumptions, would undoubtedly reveal both important differences and surprising compatibilities of the theories. But in the absence of such an analysis we cannot, without glossing over real differences, say that a particular finding helps to bring the two theories together. To do so, it would need to be shown (a) that the results demand a theory which reconciles actual differences between the two approaches; or (b) if the two psychologies are in agreement in the area in question (which would require demonstration), that the findings can be handled in terms which are compatible with both.

To say, in other words, that both personal factors and reality factors determine a perceptual effect is to pose a problem. It is to point out that we need a human psychology which will include both kinds of factors; but it is not to say that such a psychology will be compatible with Gestalt psychology or with psychoanalysis or both. To the present writer it seems more likely that a finding which cannot be handled adequately within the framework of either of these existing systems will demand, not a reconciliation of the two admittedly incomplete theories, but rather a



new theory. This point will be discussed below.

A final example, which will show again how eclectic reconciliation of differences may be achieved at the expense of a specific theory in the area of the controversy, may be taken from Welch (17). This author has offered some fundamental propositions which he believed should be acceptable both to Gestalt psychology and to contemporary behaviorism. For example, "perceiving is the result of a stimulus compound producing effects upon the sense organs which establish brain traces similar to or in otherwise related to brain traces formerly established" (17, p. 181). Surely everyone will agree that present percepts are related to traces of past ones; where dispute exists it concerns the specific effects of past experience on perception. Welch has succeeded in reconciling the differences by omitting the specific area of controversy. Again, we are told:

In interacting with its environment, the organism changes in many ways. . . . [Among other changes] it may learn. Learning is the effect of a stimulus compound or stimulus compounds upon the nervous system of the organism and the responses which these evoke, that makes possible the establishment of new responses, as a result of such experiences (17, p. 187).

Many psychologists may, indeed, accept this as a rough definition of learning. But when one leaves this level of generality and raises the question of how this process is to be envisaged, this happy harmony disappears. Hilgard points out: "There are no laws of learning which can be taught with confidence" (12, p. 457). Likewise no one questions the fact that "behavior of any type is the result of the interaction of the organism and its environment" (17, p. 176). But what is the nature of this interaction? What are the roles of organism and environment? Here are

questions on which different writers have taken divergent positions. (Cf. 11.) While no one will disagree that memorizing and generalizing (17, pp. 181, 182) occur, and while Welch's definitions might provoke little controversy as rough identifications of the phenomena in question, different theories exist about the nature of these processes.

It will be clear from the above discussion that the existence of facts which all psychologists accept is irrelevant to the problem of eclecticism. Likewise the circumstance that some developments in psychology have called attention to facts ignored by others has no bearing on the issue. The important questions are: How are these facts understood? and What is their place in the overall theoretical system? Even where agreement exists as to the facts, differences are current with respect to these questions.

In all the examples considered here, it would appear that differences have been reconciled and controversy eliminated at the price of obscuring the issues with which research is concerned in contemporary psychology.

Boring, years ago, pointed out the productive role of controversy in scientific research (2). Not only does the eclectic lose prematurely the advantages of controversy, he may to some extent give up the advantages of theory as well. The above discussion contains the suggestion that the eclectic at times renounces specific theory in the area of a controversy in order to reconcile differences. This statement will be qualified below. But now attention must be drawn to a consequence of the intimate relation between fact and theory.

There is a certain amount of fact that can be discovered in the absence of any theory. For example, time errors forced themselves to the attention of psychologists who were concerned with quite different problems. For the most part,



however, problems for investigation arise out of the theories one holds. New facts are discovered in the course of research designed to test one's hypotheses. To the extent, therefore, that the eclectic gives up specific theory in the area of a controversy, he is handicapped in the discovery of new facts.

Closer examination will, however, often show implicit theories which may contradict the eclectic's avowed intention by placing him in a position on one side or the other of the (now only implicit) controversy. Woodworth, for example, deals with transfer in terms of *carrying over* something from one learning situation to another rather than in terms of *application* of what has been learned to the training tasks as well as to the new ones. That is, transfer is seen as occurring because knowledge acquired in the original training is carried over to the new activity; the new tasks, to the extent to which they are similar to the learned ones, are considered already partially learned. The alternative is ignored that what is learned is not tasks but principles or other whole properties; thus the training activity may merely provide examples of the use of the principle which can be applied equally to the new situation. Woodworth's theory is thus close to a theory of identical elements in this respect, opposing one derived from the study of learning by understanding (13, Chap. 5) and, indeed, unable to deal with many cases of such learning.<sup>7</sup> This

<sup>7</sup> If a theory of common factors were correct, there should never be more than 100 per cent transfer, since two activities cannot have more than 100 per cent of their factors in common. Yet, as Katona has shown (13), cases exist in which performance on the test activity is superior to that on the training task.

It is of interest to note also that Woodworth's theory, while it succeeds in reconciling the differences, appears to lump together cases that do not belong together. There is evidence that transfer of specific data is dif-

ference is particularly impressive since, as will be illustrated immediately below, Woodworth is by no means opposed to learning by understanding; it is another instance of the confusions which eclecticism breeds.

In another place Woodworth calls attention to the following controversy:

Among present-day theories of learning those which emphasize reinforcement or the law of effect minimize the perceptual factor, often stigmatizing it as "mentalistic" and impossible to conceive in physical terms, while those which emphasize perceptual learning are apt to deny any direct importance to the factor of reinforcement (20, p. 119).

In attempting to show that there is "no obvious incompatibility" of these two factors, he makes (explicit) assumptions about learning as a cognitive process<sup>8</sup> which would be likely to be unacceptable to many S-R theorists, and (both implicit and explicit) empiristic assumptions about perception<sup>9</sup> which many

ferent, in process as well as in the magnitude of the effect, from the application of principles derived from one set of data to new material. (Cf. 13.)

<sup>8</sup> For example: "As to connections, several may be established before the conditioning is complete, but the primary one connects the conditioned stimulus with the meaningful character it acquires as the first event in a regular sequence" (20, pp. 121-122). Also "In experiments that offer alternatives and demand a choice, what has to be learned is a distinction between stimulus-objects and not between motor responses. . . . What has to be learned is the difference between the two alleys" (20, p. 122).

<sup>9</sup> "When a new percept is in the making—when an obscure stimulus-complex is being deciphered, or when the meaning of a cue or sign is being discovered—an elementary two-phase process is observable. It is a trial-and-check, trial-and-check process. The trial phase is a tentative reading of the sign, a tentative decipherment of the puzzle, a tentative characterization of the object; and the check phase is an acceptance or rejection, a positive or negative reinforcement of the tentative perception" (20, p. 124).

Among the implicit assumptions seems to be the view that organization is not primary in

cognitive theorists might find equally unacceptable (assumptions, incidentally, which are not necessarily consistent with those about the learning process).<sup>10</sup>

Several questions suggest themselves with respect to the theory implicit in eclectic solutions. (a) A question worth examining is whether there is a tendency for such implicit theory to be too heavily weighted in the direction of traditional theory. As the above examples show, this need not always be the case; but it seems plausible to think that when theory is not explicit, and thus not examined, it draws upon doctrines prevailing both in psychology and in the culture in general rather than upon the newer and less widely accepted theoretical currents. In a similar connection Köhler has pointed to a certain conservatism in eclecticism:

... it has been said with approval that psychology now tends to be eclectic. Again, we have been told that in psychology we had better stay in the middle of the road. I cannot agree with these prescriptions because, if they were followed, psychologists would have to look first of all backward. In an eclectic attitude, they would be too much concerned with ideas which are already available; and, in attempting to find the middle of the road in psychology, they would have to give too much attention to the tracks along which others have moved before them. Such attitudes could perhaps be recommended if, in

perception, nor prior to the effects of learning; as well as the idea that there is no fruitful distinction to be made between perception and interpretation.

<sup>10</sup> As a final illustration, Welch states: "This distinction between elementary and higher forms of learning involves the distinction between a situation where the new elements are simple in nature, or simple in character and are simply integrated, and a situation where the new elements are complex and integrated in a complex manner" (17, p. 188). This statement implies an elementaristic view of the learning process—learning being envisaged as the integration of elements—which would be far from acceptable to all the psychologists Welch is trying to reconcile.

research, security were an important issue. Actually there is no place for it in this field. In research, we have to look forward, and to take risks (15, p. 136).

(b) Another question which arises in connection with the theory underlying eclectic solutions is the following: since such theory is often implicit, and thus unexpressed and unexamined, is it adequate to lead to the discovery of new facts? For example, since the idea of "carrying over" (i.e., as opposed to that of "application") is only implicit in Woodworth's theory of identical components, it seems unlikely that it would be subjected to test. Or again, the implicit elementarism in Welch's statement about learning (cf. Footnote 10) is unlikely to be tested since the author's main focus is on other aspects of the statement.

Also worth looking into in connection with the theory implicit in eclectic solutions are the questions of its adequacy for ordering the facts and its susceptibility to proof or disproof. For example, to say that "something" is transferred is too unspecific a statement of the conditions of transfer to test empirically. Any finding of transfer seems to confirm it, and there is no result which could disprove it. Again, it has been suggested above that the theory implicit in a given eclecticism is not always internally consistent. This is a question which deserves examination in connection with particular eclectic psychologies.

We may summarize the discussion so far by saying that eclectics have to a large extent succeeded in resolving conflicts in psychology by ignoring differences and obscuring the issues. Some reasons for dissatisfaction with such solutions have been indicated.<sup>11</sup> Is there no

<sup>11</sup> It is of interest to note that eclecticism seems to have presented similar problems in other fields of knowledge in their comparative



alternative? It seems to the present writer that reconciliations can be reached in psychology only by focusing on the existing differences, examining them, and carrying on research to settle issues. If this is eclecticism, it is eclecticism after the fact rather than the prevailing eclecticism before the fact. And it is clear that it will not be a matter of reconciling existing theories. Since competing theories on any particular issue in psychology today—or competing psychological systems—each tend to be plausible and to be supported by evidence, it is unlikely that any one will win a clear victory over the others. Yet none can offer a fully satisfactory explanation—or else the controversy would not exist. Controversies do not exist in science with regard to processes which are fully understood. Thus the task seems to be one of arriving at new, more comprehensive theories of the processes in question.

youth. I quote an observation on the medical science of a century ago: ". . . And as the rules derived from fundamental truths seemed to come into unsolvable contradiction with the experiences and the sanctioned standards of practice, there sprang up under the name Eclectic the representatives of sober elucidation, of the *juste milieu*, of the medium of the extremes. The breach between theory and practice, which they feared, was avoided or postponed if theory gave up the pretension to penetrate into particulars and if practice agreed that, because of its youthful immaturity, it should be excluded from counsel, and progress in silence and in hope. The conflict was settled and peace was achieved, not by the reconciliation of the parties, but by separating them. The so-called impartial examination of the facts should lead only to a middle road between them. [The eclectics] thought they had principles and avoided their application; they proclaimed themselves free and in practice clung to the consequences of old dogmas. They practiced tolerance not because they included the truth of each dogma, but because a chasm existed between theory and life, beyond which theory didn't matter" (9, p. 9).

An example should make this clear. It seems safe to say that theories of forgetting arising out of experimental psychology have found no adequate place for the facts of repression. Nor have the psychoanalysts succeeded (or tried) to bring these facts into relation with a general theory of memory and forgetting. Can the two kinds of theories be brought together? It seems to me that the most fruitful starting point is not the attempt to reconcile existing theories. Actually, useful theories of repression do not exist. (Cf. 6 for a similar point, more generally stated.) It is hardly sufficient to say:

Repression proceeds from the ego, which, possibly at the command of the superego, does not wish to be a party to an instinct cathexis originating in the id. Through repression the ego accomplishes the exclusion from consciousness of the idea which was the carrier of the unwelcome impulse (7, p. 19).

This statement contains no hypothesis about the processes involved, about how repression can possibly be brought about. Thus there seems to be no point to attempt to reconcile the theories of experimental psychology and of psychoanalysis on repression; neither has an effective theory in this area. What we need is to look into the processes themselves, in the light of what we know about forgetting in general (cf. [10]). Can affective processes act, for example, to produce a failure of the Höfding function—i.e., that selective interaction between present process and memory trace which is the basis of recognition and the first step in the process of recall? (Cf. 14, pp. 126 ff.) Under what conditions can emotional and motivational processes introduce interferences? Answers to such questions might lead not only to a hypothesis about the nature of repression, but might also introduce considerable modification into our



present theories of the nature of forgetting in general.

The eclectics are, of course, right in maintaining that where a genuine controversy exists in psychology, and where evidence seems to support both sides, there is likely to be some truth to both positions. But they solve their problem too soon. Existing theories cannot be made more comprehensive by adding divergent ones together. They can be broadened to include all the relevant evidence only by looking more deeply into the phenomena with which they are concerned; and this means arriving at new theories.

At this point the parallel between productive solutions of theoretical problems and of personal problems becomes striking. In connection with the reconciliation of opposites within the personality, C. G. Jung points out that conflicts are never resolved on their own level. They are outgrown. Only on a higher level can you see both sides.

#### SUMMARY

Examples have been presented to show that eclectics tend to resolve conflicts in psychology by glossing over real differences and obscuring the issues. Such solutions achieve harmony at the price of specific theory in the area of the controversy, and thus sacrifice fruitfulness in the discovery of new fact. Closer examination often reveals implicit theories underlying such solutions, but unexpressed and unexamined theory can hardly be expected to equal explicit hypotheses either in fruitfulness or in adequacy in dealing with known facts.

It is here suggested that differences need to be resolved in psychology not by denying them and attempting to combine existing theories, but by focusing on the differences and using them to get a better view of the relevant phenomena. We will achieve more compre-

hensive theories not by combining existing ones but by understanding better the processes in question.

#### REFERENCES

1. ABT, L. E. A theory of projective psychology. In L. E. Abt & L. Bellak (Eds.), *Projective psychology*. New York: Knopf, 1950.
2. BORING, E. G. The psychology of controversy. *Psychol. Rev.*, 1929, 36, 97-121.
3. BORING, E. G. *A history of experimental psychology* (2nd ed.). New York: Appleton-Century-Crofts, 1950.
4. BORING, E. G. The nature of psychology. In E. G. Boring, H. S. Langfeld, & H. P. Weld (Eds.), *Foundations of psychology*. New York: Wiley, 1948.
5. BRONFENBRENNER, U. Toward an integrated theory of personality. In R. R. Blake & G. V. Ramsey (Eds.), *Perception, an approach to personality*. New York: Ronald Press, 1951.
6. BRUNER, J. S. Freud and the image of man. *Amer. Psychologist*, 1956, 11, 463-466.
7. FREUD, S. *The problem of anxiety*. New York: Norton, 1936.
8. FREUD, S. *Beyond the pleasure principle*. London: Hogarth Press and the Institute of Psycho-Analysis, 1950.
9. HENLE, J. *Handbuch der rationalen Psychologie*. Erster Band. (2nd ed.). Braunschweig: F. Vieweg u. Sohn, 1846.
10. HENLE, MARY. Some effects of motivational processes on cognition. *Psychol. Rev.*, 1955, 62, 423-432.
11. HENLE, MARY. On field forces. *J. Psychol.*, 1957, 43, 239-249.
12. HILGARD, E. R. *Theories of learning*. (2nd ed.) New York: Appleton-Century-Crofts, 1956.
13. KATONA, G. *Organizing and memorizing*. New York: Columbia Univer. Press, 1940.
14. KÖHLER, W. *Dynamics in psychology*. New York: Liveright, 1940.
15. KÖHLER, W. The scientists and their new environment. In W. R. Crawford (Ed.), *The cultural migration*. Philadelphia: Univer. of Pennsylvania Press, 1953.
16. LEWIN, K. *Principles of topological psychology*. New York: McGraw-Hill, 1936.
17. WELCH, L. An integration of some fundamental principles of modern behavior-

- ism and Gestalt psychology. *J. gen. Psychol.*, 1948, 39, 175-190.
18. WITKIN, H. A., LEWIS, H. B., HERTZMAN, M., MACHOVER, K., MEISSNER, P. B., & WAPNER, S. *Personality through perception*. New York: Harper, 1954.
19. WOODWORTH, R. S. *Experimental psychology*. New York: Holt, 1938.
20. WOODWORTH, R. S. Reinforcement of perception. *Amer. J. Psychol.*, 1947, 60, 119-124.
21. WOODWORTH, R. S. *Contemporary schools of psychology*. (Rev. ed.) New York: Ronald Press, 1948.
22. WOODWORTH, R. S., & MARQUIS, D. G. *Psychology* (5th ed.). New York: Holt, 1947.
23. WOODWORTH, R. S., & SCHLOSBERG, H. *Experimental psychology*. (Rev. ed.) New York: Holt, 1954.

(Received January 9, 1957)

## VISUAL PERCEPTION: AN EVENT OVER TIME

GUDMUND SMITH<sup>1</sup>

*Research Center for Mental Health, New York University*

Most psychologists would certainly agree that perception—or more broadly, cognition—like other aspects of behavior, is not an instantaneous datum but, instead, that it should be considered an event over time. There are only a few of them, however, who have consistently entertained this self-evident notion when designing their experiments and their explanatory models. Until recently only the Leipzig psychologists, in their study of the *Aktualgenese* of percepts, had tried to make the microgenesis of perception the center of their experimental and theoretical approach. Their admitted failure to call attention to the genetic point of view must in part be attributed to their method of developing concepts which were mainly of a descriptive order (for further criticism, also methodological, see 8), and in part to the inability of current schools of psychology to include the Leipzig findings in their explanatory systems.

This paper is not intended as an attempt to rehabilitate the Leipzig school, however justified that may be, but an attempt to reconsider the genetic approach to problems of visual perception. In view of recent contributions within the fields of personality and perception, such an approach may seem more rewarding than it did thirty years ago; Hebb, for instance, in his physiological theory, takes into account the "phase sequence" as a prerequisite for the organization of a percept (5). The present author and his collaborators have tried to demonstrate, more specifically, that "serial" interpretations of perceptual reports continuously repeated in the same experimental situation (a situation which

should be so new or so complicated as not to allow an adequate report after only one trial) often tended to reveal more about an individual's adaptive mechanisms than interpretations only of the final, stabilized reports or of the usual summary scores (10, 13). Our basic consideration thus implies that we look upon perception (cognition) as a process of organization, emphasizing that it needs time to be prepared, evolved, and established. This consideration has been developed in earlier papers (e.g., 11) and partly in cooperation with Kragh (8), although some points where the present author differs from him will be mentioned later (cf. 12).

### BASIC ASSUMPTIONS

The process of visual perception should not be confused with what is generally called percepts, which are products of the process fitted into a frame of reference of outside reality. It would be even more correct to say that percepts normally are the products of late phases in a process of organization, the preparatory phases of which are therefore valid objects of study. Some important characteristics of these early or preparatory phases are implied already in the basic consideration. First of all, we have to assume that these stages do not generally produce percepts; only rarely do we become aware of our own perceptual processes. This assumption is of extreme importance for any genetic theory of perception, since it indicates that we do not have to conceptualize preparatory stages in terms of finished products, e.g., as faint copies of established percepts or of stimuli, or as reflections of physiologi-

<sup>1</sup> Now at the University of Lund, Sweden.



cal or other processes hypothesized on the basis of end-product behavior (11). It is more natural to assume that preparatory stages often differ from end stages just because they have not been able to produce percepts. Such premature products of early stages would probably be suppressed, at any rate, in favor of something more suitable for the frame of reference of outside reality.

When the perceptual process is inhibited, as in a tachistoscopic experiment where exposure times are too brief for a correct perception of the stimulus, early stages of the perceptual process can be forced to produce a percept. But we know that these percepts are often diffuse and equivocal. Another characteristic of preparatory stages might therefore be that they embrace numerous "possibilities" for further development. Let us, to simplify matters, call it an unspecific attention toward something that is going to emerge (e.g., to be seen), a mobilization of a system for many possible activities. As the ideal perceptual process continues in a situation with prolonged stimulation, however, more and more of these possibilities fall into the background in favor of the one possibility that coincides with the stimulus (or, with a normative conception of the stimulus), in favor of the unequivocal reality of the individual (or his culture, or his language). Only fairly unequivocal stages can produce percepts, at least in so-called normal human beings. We should then state, more generally, that *one* possibility has to be singled out for the perceptual process to produce a percept—and, of course, not always the possibility that is correct from the point of view of stimulus. We know well by now that even "normal" people can perceive incorrectly.

Now we can be a little more specific about some additional characteristics of the preliminary stages. It seems reasonable to assume that it is not im-

possible for them to harbor completely contradictory possibilities of development, i.e., contradictory from the point of view of outside reality (stimulus). When reversible figures are presented near or below the threshold of awareness, many individuals seem to recall both "sides" of them at the same time;<sup>2</sup> or they report the correct percept as well as various incorrect ones, the full-fledged Gestalt along with fragments of it, etc. Although we conceive preliminary phases as encompassing contradictory possibilities "side by side," we should not think of them as a number of superimposed percepts. As stated above, the preparatory stages should not be conceptualized in terms of established ones. It would probably be more fruitful to try to relate preparatory phases to final ones in the same formal way as primary processes are related to secondary ones in psychoanalysis; i.e., only the final phases have to adhere to the logic and organization of outside reality. A later stage, however, cannot be fully understood with reference only to this outside reality, to the stimulus. The main emphasis, from the genetic point of view, is upon the sequence over time, the interpretation of late stages in the light of earlier ones. Even if the later stage is different in scope and organization from previous stages and, emerging on top of the sequence, is in control of them, the roots of this stage are still to be found in the preparatory part of the perceptual process.

#### THE AKTUALGENESE APPROACH

The Leipzig school made an attempt to capture the perceptual process by presenting the same stimulus, first at very brief exposures, then at gradually prolonged ones. Alternative methods were also employed, e.g., when the intensity

<sup>2</sup> Charles Fisher: Lecture at New York University, 1956.

of the stimulus was increased from one trial to another. It was assumed that the shorter the presentation, or the lower the intensity of the stimulus, the earlier the stage at which the process was forced to produce a percept. Kragh (8) has refined the tachistoscopic method considerably: he used new types of stimulus material, e.g., TAT-like pictures, and tried to define the "fractionized" stimulus more rigorously than before. But it is even more important that he developed his serial or micro-genetic method in close relation to a conceptual model of perception-personality. The two concepts supporting his model are *construction* and *reconstruction*. Construction refers to the evolving process of organization, the ideal end product of which is the correct percept. But our knowledge about such a process is always based upon a series of reports of gradually prolonged stimulation. In order to describe the temporal organization of a percept we thus use the method of reconstructing stages before the final stage; i.e., various phases of the process of construction have been enticed prematurely to produce percepts fitted to the conceptual level of the end product. Nevertheless, a series of reports of a stimulus exposed for gradually extended times should, according to Kragh, reflect the microscopically short genesis of a percept even if the filtering process of reconstruction deprives the early stages of some of their qualities. Since such a series of preliminary reports implies a continuous progress toward the final percept, however, qualities lacking in early stages may frequently be inferred from later stages.

It is not our intention to question here the assumption that these perceptual reports belong to the *same* process of organization. But we are less optimistic than Kragh as to the possibility that a report in his experiment can be regarded

as a veridical representation of the early stage at which the perceptual process was supposed to have stopped. Kragh's concept of reconstruction pays attention to the fact that, in *Aktualgenese* experiments, early stages have to be communicated on an end-stage level, but it does not refer explicitly to the equally important fact that, since percepts have to fit into reality, a number of other perceptual possibilities must have been disguised or omitted in order to facilitate the event of a percept. Thus, in *Aktualgenese* experiments, a series of early reports are likely to reflect the early part of the perceptual process very insufficiently, or at least in the same complicated manner as dream condensations reflect primary as well as secondary processes. But this does not preclude the assumption that they represent very powerful possibilities of development which may manifest themselves in one experiment after the other, only slightly disguised. One possibility of this kind would naturally be the correct percept, as was demonstrated by the compulsive group in Kragh's experiments (8); these people tended to proceed directly from blank stages to correct ones, i.e., to avoid or isolate all incorrect possibilities. In other subjects it was evident that many of the alternative perceptual possibilities were related to important personal experiences far back in life.

Another important assumption is that the perceptual process developed over an intermittent series of short presentations is equivalent to the process accompanying one long presentation of the same stimulus. Kragh maintains that there is an approximate similarity between everyday perception and perception in *Aktualgenese* experiments, partly because our fixation even of immovable objects is likely to be intermittent. But these intermittent fixations in normal perception are not accompanied by conscious



perceptual products differing in degree of clarity and correctness; i.e., while tachistoscopic experiments tend to fix the perceptual process at various stages of organization, the normal process is probably more continuous. It seems only reasonable to assume that the course of perceptual organization will change more or less drastically when one of its erroneous alternatives appears as a conscious percept. The *Aktualgenese* method also tends to hamper the feedback mechanics which probably play an important role in normal perception (1). These and previous considerations regarding the influence exerted on the perceptual process by the *Aktualgenese* method may be important for the assessment of conclusions based on *Aktualgenese* data. But they give us no reason to abandon the working hypothesis that perception is a process of organization extending over microtime.

#### ELABORATION OF SOME BASIC ASSUMPTIONS

One of the basic assumptions for this model of perception is that the "germ" of a correct percept is embodied already in early stages of the perceptual process. Naturally, this would be true only in cases where, from the stimulus point of view, a correct perception was possible, i.e., where the retina was able to register a sufficiently large part of the object, etc. If this was not the case, the perceptual process would have to be reinforced by way of a new fixation. Any repeated or prolonged fixation would certainly strengthen the correct possibility. Let us assume, at the same time, that early stages are more prone to produce ideas, dreams, etc., than percepts—an assumption in complete accordance with previous assumptions about these stages. It is hardly surprising, then, that in dreams which their patients reported on the morning after

the experiment, Poetzl (9) and Fisher (4) found exact copies of those parts of the stimulus that had not been consciously perceived together with fragments of them. We are convinced, furthermore, that many more "distorted" products of the same process would be found in these dreams if there was any acceptable method to account for them. The more we reduce the intensity or duration of stimulus the easier it will be for "incorrect" products to appear instead of the correct one. And when the stimulus itself is equivocal, as in many projective tests, there is no longer any "correct" possibility.

The exact nature of the perceptual process hypothesized here may still seem a little mysterious. The primary source of knowledge is perceptual reports, either reports of end stages obtained under optimal stimulus conditions or of such quasi end stages as imply that the perceptual process has been prematurely interrupted. An alternative to the latter methods will be described later in this paper. But these sources are still insufficient to account for the perceptual process. If our main postulates are correct we can learn much about it, especially its early phases, by studying such products as dreams, images, etc., i.e., products that cannot be accepted as outside reality here and now. In this way we bring the study of perception into the broader field of personality. We may even say that the process of perceptual organization is a reflection of personality in the same way as its end products are reflections of the reality built up around the individual. When we thus postulate that we can learn much about basic processes in the individual by reconstructing the construction process of perception, we have not implied anything mysterious or even radically different from current conceptions about perception-personality.



## PERCEPTION-PERSONALITY

The process of perceptual organization may be described essentially in formal terms or in terms mainly of content. It is expected to show characteristics typical of the individual, so-called consistencies or structures, and also characteristics closely related to the current situation. Some structures would be more typical of the early phases of the process than of later ones because the objectivized conception of reality is similar in most individuals; this possibility has been realized in much present-day perception-personality research where the intensity or clarity of the stimulus is cut back. But this would not be enough to account for the structures reflected in the perceptual process. Structures also refer to the longitudinal aspects of the process, the way in which certain end stages are reached and others warded off, to all those aspects of control of the organization process which are termed cognitive attitudes or system principles (2, 6). For instance, while the perceptual processes of compulsive people seem to be controlled by an all-or-nothing principle, by avoidance of all stages between a blank report of stimulus and a correct one, those of other people tend to develop more continuously over a series of more or less incorrect (subjective) reports (8, 15). However, we do not have to explain these and other characteristics of the process by referring to constructs outside the process, or method, itself. The explanatory constructs are given in the process of construction and the operation of reconstruction by which this process is unveiled.

It would be even more correct to say that the explanatory constructs of any temporarily limited process of organization are ultimately given in that all-embracing continuous process of construction the qualities of which we find

in the life history of the individual. A perceptual process is naturally to be regarded as a continuation of this chain of adaptive events. Then, in its various stages the perceptual process must also be marked by primary or secondary characteristics of personality organization; the incorrect "possibilities," for instance, are not accidental (how could they be?), but often reflect those experiences out of which the current individual was shaped (8). Thus perception has the same roots as other forms of behavior, although its end products manifest themselves in the specific frame of reference of outside reality. We may go even further and assert that percepts and some aspects of simultaneous overt behavior are two sides to the same event, an assumption implied in our common acceptance of verbal reports as indicative of percepts. But while numerous unsuitable products of a process of organization cannot manifest themselves as percepts, they may appear in other forms of behavior which do not have to be so closely fitted into a normalized framework. And these reflections of preparatory stages, as said above, are most obvious in such products as were discarded in the perceptual process as unsuitable representatives of outside reality, viz., our own dreams, images, hopes, phantasies and the like (see also below concerning the measurement of nonverbal forms of behavior). Only by studying all behavior manifestations can we learn the full story of perception.

## PERCEPTION BEYOND AWARENESS

The problem of perception beyond awareness must be considered crucial for the approach outlined here, which implies that an exhaustive study of the problems of perception cannot be restricted to the characteristics of conscious end products. In collaboration with a number of other psychologists, the present author has tried a new tech-

nique for investigating preconscious aspects of the perceptual process. This technique has been reported in detail elsewhere (7, 14, 15, 16). The following procedure may be regarded as typical (cf. also 3, 17). Two stimuli are exposed in rapid succession in a tachistoscope. The exposure time for the second of these stimuli (*B*) is so long as to admit of a correct, stabilized perception of it, while the first stimulus (*A*) is presented too briefly to be conceived as an independent entity by the subject. Since *A* and *B* appear on the same part of the stimulus field, one can say that the perception of *A* is extinguished by the perception of *B*. But in a number of subjects the subliminal perception of *A* affects the perception of *B* in various ways; the two percepts seem to coalesce into a changed perception of *B*. This is to say, we study a perceptual process (*A*) beyond awareness by reading off its effects on an adjacent, conscious percept.

In some of our studies, the phase of the *A* process influencing the perception of *B* seemed to have been comparatively late, at least so late that the stimulus "possibility" asserted itself beyond other possibilities. This was evident in an experiment where a fan-shaped line pattern (Stimulus *A*) influenced a square (Stimulus *B*) as if the former subliminal stimulus had been correctly perceived; i.e., the square developed into the same rhomboid as in the well-known illusion with *B* on the lined background of *A*. But although the *A* process thus changed the reports of *B* in a direction predictable from the correct *A*, it was still too weak and diffuse to assert itself as a full-fledged percept beside the *B* percept. If it should be represented at all it had to be represented within the context given by the latter percept. However, while this experiment was designed to prove only that the perception of *B* may be affected by a perceptual process beyond

awareness, later experiments indicated that the relations between the preconscious process and associated conscious reports differed interindividually (15, 16). Here stimuli were more complex and were made up of meaningful line drawings. In other words, the characteristics of early stages of the perceptual process, as implied in reactions to *B*, did not reflect only characteristics of the subliminal stimulus but also reflected basic cognitive attitudes, e.g., the defence mechanisms of projection and isolation. Thus in a paranoid group the subliminal influence of *A* upon the perceived reality context (*B*) was often drastic, and varied from one individual to another, whereas in a compulsive group the *A* process seemed to be isolated from *B* until *A* emerged as a correct, independent entity (15).

It seems rather self-evident that a vague and equivocal conscious percept would be more easily affected by a subliminal process than would a clear and well-structured percept. In an experiment with a complex *B*, a face, we found greater effects in the beginning of a series of reports than later, when the impression of the face became more stabilized; only when the exposure time for *A* approached threshold values did its influence on the perception of *B* increase again (16). It was interesting to note that the effect of subliminal stimulation (the words HAPPY and ANGRY) was not necessarily directly related to the conventional meaning of these words; i.e., ANGRY did not produce only angry expressions in the face but more often also tense, anxious, serious, pensive, and similar expressions. This may partly confirm our assumption that the early stages of a perceptual process include more varied "possibilities" than those related to conscious, stabilized cognition, though we also have to consider that interaction between the subliminal process and con-



scious thoughts about the face probably contributed to the modification of meaning. If Stimulus *A* is very novel, unusual, or controversial, there is reason to believe that, in many individuals, other possibilities than the correct one will influence the *B* percept, or, in other words, that the difference between sub- and supraliminal conditions will be considerable (7). The same will be true of differences between individuals in the subliminal condition.

In most of these experiments on perception beyond awareness, the actual descriptions constituted the basic core of information. But if subjects could become sufficiently relaxed, images and by-remarks accompanying the reports might be even more important, as demonstrated by Kragh (8), who used a free-association technique in his studies. *A* tendencies which cannot assert themselves within *B*, and which cannot be represented within this frame of outside reality, may very well appear in images, associations, denials—or, for that matter, they may influence other, nonverbal forms of behavior. The study of these behavioral aspects should become more crucial the more resistant to change are perceptions of *B* stimuli. In one of the studies just mentioned, for instance, differences in latency of reactions to a *B* stimulus presented in alternating combination with the subliminal stimuli ANGRY and HAPPY seemed to *increase* as differences in verbal description of *B* in these two pairings *decreased*. In other words, where the two different *A* stimuli did not produce differences in a subject's descriptions of *B*, a face, they produced differences in the time taken before he started to react verbally (16). Dreams reported the morning after the experiment have also proved important to account for the structures and meanings in the preconscious *A* process. However, only by extending these studies to still more general aspects of per-

sonality organization can we fully assess the implications of our findings. A preliminary attempt was made in the paper just referred to.

The experiments described here naturally represent only a few variations of the basic method which may be applied in many forms to a variety of perceptual phenomena. Generally, I think that in these experiments, where the *A* process does not have to manifest itself as independent conscious percepts, it will be possible to explore very important aspects of the perceptual process that cannot be reached in experiments of the *Aktualgenese* type. One of the specific problems raised by the new method concerns the interrelations between a perceptual process, the products of which hold the center of awareness, and more marginal processes continually evolving at the same time—interrelations of immediate relevance for the study of everyday perception and, as has been illustrated here, of basic mechanisms of control and adaptation. But most important of all, both types of experiments imply a challenge to the traditional assumption that perception is an instantaneous event the mechanisms of which are reflected only in conscious, unequivocal products.

#### SUMMARY

In this paper perception is considered as a microscopically short process of organization, the prestages of which are therefore important objects of study. Some theoretical and empirical approaches to this genetic analysis of perception (and personality) are discussed together with the topic of subliminal perception, which is of particular importance in this connection.

#### REFERENCES

1. ALLPORT, F. H. *Theories of perception and the concept of structure*. New York: Wiley, 1955.



2. ANGYAL, A. *Foundations for a science of personality*. New York: Commonwealth Fund, 1941.
3. CHEATAM, P. G. Visual perceptual latency as a function of stimulus brightness and contour shape. *J. exp. Psychol.*, 1952, 43, 369-380.
4. FISHER, C. Dreams and perception. The role of preconscious and primary modes of perception in dream formation. *J. Amer. Psychoanal. Ass.*, 1954, 2, 389-445.
5. HEBB, D. O. *The organization of behavior*. New York: Wiley, 1949.
6. KLEIN, G. S. Perception, motives, and personality. A clinical perspective. In J. L. McCary (Ed.), *Approaches to personality*. New York: Logos Press, 1956. Pp. 121-199.
7. KLEIN, G. S., SPENCE, D. P., HOLT, R. R., & GOUREVITCH, SUSANNAH. Cognition without awareness: subliminal and supraliminal influences upon conscious thought. *J. abnorm. soc. Psychol.*, in press.
8. KRAGEH, U. *The actual-genetic model of perception-personality: an experimental study with non-clinical and clinical groups*. Lund, Sweden: Gleerup, 1955.
9. POETZL, O. Experimentell erregte Traumbilder in ihren Beziehungen zum indirekten Sehen. *Z. Neurol. Psychiatr.*, 1917, 37, 278-349.
10. SMITH, G. Twin differences with reference to the Müller-Lyer illusion: a study in modal summaries and serial change. *Lunds Universitets Årsskrift. N.F. Avd. 1*, 1953, 50, 1.
11. SMITH, G. The place of physiological constructs in a genetic explanatory system. *Psychol. Rev.*, 1954, 61, 73-76.
12. SMITH, G. A new psychological model. *Theoria*, 1956, 12, 61-69.
13. SMITH, G., & KLEIN, G. S. Cognitive controls in serial behavior patterns. *J. Pers.*, 1953, 22, 188-213.
14. SMITH, G., & HENRIKSSON, MAJ. The effect on an established percept of a perceptual process beyond awareness. *Acta Psychol.*, 1955, 11, 346-355.
15. SMITH, G., & HENRIKSSON, MAJ. Studies in the development of a percept within various contexts of perceived reality. *Acta Psychol.*, 1956, 12, 263-281.
16. SMITH, G., SPENCE, D. P., & KLEIN, G. S. Subliminal effects of verbal stimuli. *J. Pers.*, in press.
17. WERNER, H. Studies on contour: I. Qualitative analyses. *Amer. J. Psychol.*, 1935, 47, 40-46.

(Received January 3, 1957)

## RECENCY, FREQUENCY, AND PROBABILITY IN RESPONSE PREDICTION

JOHN E. OVERALL AND W. LYNN BROWN<sup>1</sup>

*The University of Texas*

It has been ten years since Voeks (13) aroused psychological interest by showing that contemporary learning theory can be used as a basis for individual response prediction. It still remains a point of controversy as to whether *Ss* utilize all past experiences in the situation (frequency), or whether only the most recent experience determines the response (postremity). Voeks found that postremity predictions were the more accurate, even where frequency and postremity led to opposing predictions. Nevertheless, it was observed that postremity predictions were less accurate where they conflicted with frequency predictions than where the two did not conflict. This suggests that a theoretical model which bases predictions upon *both* frequency and recency might lead to even more effective response prediction.

### THEORETICAL MODELS FOR INDIVIDUAL RESPONSE PREDICTION

Contemporary learning theories may be divided on the basis of adherence to a recency principle or adherence to a frequency principle. A brief survey of several contemporary theories will make this problem explicit.

#### *Guthrie's Postremity Predictions*

One group of theorists, whose leading exponent is Guthrie, sees recency as the important determiner of behavior. Guthrie says that stimuli and responses become associated in an all-or-none

fashion by a single contiguous occurrence. There is no problem of strengthening; either the response is associated full strength with the stimulus or it is not. When the stimulus next occurs, the response made last in its presence will be repeated. Voeks (14) has formalized this in her principle of postremity. ". . . A stimulus which has accompanied or immediately preceded two or more incompatible responses is a conditioned stimulus for only the last response made while the stimulus was present" (6, p. 344). For the T-maze situation to be investigated in this paper, postremity would seem to lead to the prediction that the *S* will choose the same cul he chose last in the situation.

#### *Postremity-Reward Predictions*

Postremity yields a different type of prediction if the incentive value of reward is considered. Where the noncorrection method is employed and no reward follows a response, the very last movement made, after it is discovered that no reward is present, might be a "fractional" movement toward the opposite goal box. This would lead to the prediction that *S* will respond on the next trial by going where he went last, if he found reward, but that he will respond on the next trial by going to the opposite side, if no reward were encountered on the last trial.

#### *Predictions Based Upon Frequency of Response*

In sharp contrast to the Guthrian position are theories which say that all previous experiences in the situation enter into the response determination.

<sup>1</sup> The cost of this investigation was met in part by a grant from the Research Institute of The University of Texas, Austin, Texas.

Such theories neglect the principle of recency entirely.<sup>2</sup> Response predictions based upon frequency of response might be derived from Spence (9), who classifies maze learning as instrumental behavior. Since reward is not assumed to work backwards to strengthen associations made along the path to the reward, and since it is assumed to function through fractional anticipatory responses as an *incentive motivator* only, then it follows that increments of associative strength ( $sHr$ ) accrue simply as a function of number of occurrences. The more frequently an instrumental response is executed, the more associative increments are added to the strength of the habit. The importance of recent events in the learning sequence is not emphasized. Thus, frequency of response provides another basis for predicting individual behavior.

#### *Frequency of Reinforcement*

If reward is assumed to reinforce instrumental acts along the path according to a goal gradient, then frequency of reward may be employed as another basis for predicting individual responses in a selective learning situation.

#### *Algebraic Sum of Positive and Negative Cases*

The writings of Hull (7) provide still another model for predicting individual responses, although he pleads that the concept of oscillation will prevent individual response prediction from ever achieving high accuracy. Hull classifies learning in a T maze as a kind of trial-

and-error learning. Events at the goal box strengthen reactions made in arriving at the goal according to a diminishing function known as the "goal gradient." Each time reward follows the response, an increment is added to  $sEr$  associated with that response, and each time a response is not followed by reward an increment is added to  $sIr$  associated with the response. Hull states: "... We also tentatively assume, though without adequate evidence, that  $\Delta sIr$  follows the same law with the same constants as  $\Delta sEr$ " (7, p. 24). The effective reaction potential associated with one alternative in the T maze may be roughly estimated as  $sEr = sEr + sIr$ , where increments to  $sEr$  and  $sIr$  are scaled directly in terms of reinforced and nonreinforced outcomes. This procedure admittedly neglects "generalization" of both  $sEr$  and  $sIr$ , as well as the postulated progressive reduction in increment magnitudes as  $sEr$  and  $sIr$  approach their respective maximums. The defense for this procedure is that in a random 50-50 T-maze situation these factors should very largely cancel out, with generalization roughly equal for the two alternatives and with the reduction in increment magnitudes for  $sEr$  matched by a similar reduction for  $sIr$ . Consequently, we have still another frequency-type model for predicting individual responses—i.e., the algebraic sum of positive and negative instances associated with each response alternative.

#### *Probability of Reward*

Another theoretical model which yields individual response predictions is a cognitive-probability model. Tolman and Brunswik (12) conceived of the organism as relating *signs* to *means-objects*, which in turn are "causally" related to *desired goals*. However, because the causal texture of the environment is not univocal, the organism develops hypotheses based upon the probabilities of

<sup>2</sup> It should be noted that leading frequency theorists attempt to sidestep the problem of individual response prediction by pointing to lack of control and inadequacy in experimental techniques. It cannot be denied, however, that learning is a phenomenon which takes place within the organism; consequently, the final test of a theory of learning must be its ability to predict and make understandable the specific behavior of individuals.



*means-objects* resulting in *desired goals*. That route or behavior is chosen which has the highest probability of leading to reward. By considering the relative frequency of rewarded responses to total responses

$$\left( p = \frac{\Sigma R}{\Sigma R + \Sigma N} \right)$$

for each T-maze alternative, it is possible to calculate mathematically a probability estimate associated with each alternative. If this is done before each individual response, it may be predicted that *S* will choose the alternative with the greater probability of reward.

#### *Recency-Weighted Probability Model*

As outlined above, the probability model considers the recency principle no more than does a frequency model. However, it is tentatively suggested that the probability model might be revised in a manner compatible with the general theory of Tolman (11), especially to include capacities of the organism such as "retentivity," which Tolman emphasizes in his "capacity laws." This revised expectancy theory should provide the basis for more accurate response predictions which are based upon a weighting of *both frequency and recency*. It is recognized that the recency principle might be incorporated into contemporary frequency theory as easily as into the cognitive model developed here.

Three steps in the determination of a single selective learning response will be considered. First, the initial nature and magnitude of memory traces must be established. Past events in the learning sequence can affect present responses only through traces left by them within the nervous system of the organism. Secondly, traces within the memory system are not permanent; consequently their decay with time must be considered. Finally, selection of the response will be considered.

*Creation of initial traces within the memory system.* The probabilities of reward and nonreward upon which responses are based must be determined by past experiences of reward and nonreward in the situation. Since past experiences are themselves no longer present, it becomes important to understand the process whereby *traces* of past events are created within the memory system of the organism. Traces within the memory system are conceived to be positive or negative and to have greater or smaller magnitudes as a function of "expectancy."

It is proposed that the initial trace created within the memory system by a behavior outcome will be *positive* if that outcome is more "desirable" than the expected outcome, and will be *negative* if the outcome is less "desirable" than the expected outcome. An outcome which is exactly equal to the expected outcome will be neither positive nor negative. For example, Tinklepaugh (10) found that a lettuce leaf which would have had a positive valence under certain other conditions was rejected as negative by his chimps when they expected to find a banana. Crespi (2, 3) obtained different performance values as a function of expectancy in partial and in 100 per cent reinforcement situations, and Elliott (4) demonstrated shifts in maze performance of rats resulting from changes in the type of food reward.

It is proposed that the initial *magnitude* of a positive or negative memory trace is a function of expectancy. The farther a behavior outcome departs from the expected outcome, the greater the magnitude of the trace created by it within the memory system. Thus, if the expectancy of reward is low, a single rewarding outcome will create a trace of relatively large magnitude. In a situation of low-reward expectancy a single nonrewarding outcome will create a trace

of relatively small magnitude. Conversely, if the expectancy of reward is high, then a single rewarding outcome will create a relatively small trace and a single nonrewarding outcome will create a relatively large trace.

*Decay of traces within the memory system.* Probabilities of reward cannot be calculated directly from the initial trace magnitudes because earlier traces have decayed more than recent ones. For the purposes of this paper no theory of decay is necessary; however, it is tentatively assumed that simple decay with time adds to the effects of interference from new learning to produce the phenomenon. Where both time intervals between trials and opportunities for new learning remain constant in a learning situation, it is proposed that each trace within the memory system decays by a constant fraction upon the occurrence of each new trace. Suppose the constant fraction were *arbitrarily* chosen as 10 per cent. A single memory trace of initial magnitude 100 might then decay, successively, from 100 to 90, to 81, to 73, to 66, to 59, etc. These new values represent the *residual magnitudes* remaining after successive reductions of initial trace magnitude through decay in the memory system. It follows from consideration of the decay function that recent events are weighted more heavily than less recent ones in the determination of probabilities upon which responses are based.

*Selection of response.* The organism weighs the "probabilities" of reward associated with each alternative and selects the alternative associated with the greatest "probability" of reward. As stated above, it seems unlikely that the calculus of a frequency theory of probability is used. It has been proposed that decay of traces within the memory system causes recent events to have greater weight than less recent ones in determining the probability estimates

upon which responses are based. An approximation of the "subjective" probabilities which determine the organism's choice of response can be derived from the *residual trace magnitudes* discussed above. Three steps are involved. (a) An initial magnitude must be assigned to each behavior outcome. (b) A decay coefficient must be selected so that each memory trace may be reduced in magnitude upon the occurrence of each new trace. This determines residual trace magnitudes. (c) Positive *residual* magnitudes associated with a response are then summed to yield  $\Sigma Pr$ . Negative *residual* magnitudes are summed to yield  $\Sigma Nr$ . The probability of positive outcome associated with response ( $r$ ) is then calculated:

$$p = \frac{\Sigma Pr}{\Sigma Pr + \Sigma Nr}$$

This model may be used for individual response prediction by calculating the probabilities associated with each alternative and then predicting that  $S$  will select the alternative associated with the greatest probability of reward. Predictions based upon this model will thus be determined jointly by frequency and recency of reward.

An illustration will make clear the differences between predictions based upon a recency-probability model and predictions based upon frequency or traditional probability calculations. Suppose, to take an extreme example, that  $S$  is rewarded on five successive trials and that he is not rewarded for the next five successive trials. For a second example, suppose that the first five trials are nonrewarded, but that a second five trials are rewarded. Frequency of reward, frequency of response, or an algebraic summation of positive and negative outcomes fails to distinguish between the first and second examples. Likewise, probability based upon relative frequency of reward is the same for the



two examples. The recency-weighted probability model, which has been tentatively proposed here, leads to very different "probabilities" in the two examples. To greatly simplify computation, suppose that the initial trace magnitude of each of the 10 outcomes was 100, and that a decay coefficient of 10 per cent is accepted. In the first example, where reward followed only the first five trials, the sum of positive residual traces is  $\Sigma Pr = 242$ , and the sum of negative residual traces is  $\Sigma Nr = 410$ . The recency-weighted probability

$$p = \frac{\Sigma Pr}{\Sigma Pr + \Sigma Nr} = 0.37$$

In the second example only the last five of the 10 responses resulted in reward. In this case  $\Sigma Pr = 410$  and  $\Sigma Nr = 242$ . The recency-weighted probability in the second example is

$$p = \frac{\Sigma Pr}{\Sigma Pr + \Sigma Nr} = 0.63$$

The recurrence of the response is more likely where the last five of ten responses were rewarded than where only the first five of ten were rewarded.

It should be pointed out that the stochastic models of Estes and Burke (5) and Bush and Mosteller (1) incorporate several of the principles emphasized in this paper. Estes and Burke quantify association theory in such a manner that predictions derived from their model are based upon both recency and frequency. Also, in common with the cognitive formulation they state, intuitively, that a reinforced trial following a sequence of nonreinforced trials adds a larger associative increment than one reinforced trial in a sequence of reinforced trials. This function has been assigned to expectancy-based differences in initial trace magnitudes in the cognitive probability model. The Bush and Mosteller model is similar in many respects to that of

Estes and Burke; it differs from the latter model primarily in the functions assigned to reinforcement and inhibition. It appears likely that both models could lead to accurate response predictions in the experimental situations discussed below. Neither of these theories was utilized as the basis for individual response predictions in the following experiment because each requires the use of an arbitrary (until empirically established) weighting coefficient. If any invidious comparisons were to result from predictions the present writers might derive from these models, it is quite probable that selection of other coefficients would substantially change the results. Further extension of these models by their authors into the area of individual response predictions is eagerly awaited.

#### CHOICE OF EXPERIMENTAL SITUATIONS

In a great many situations employed in typical learning experiments, it is impossible to assess the importance of *order* of events in the learning sequence. It may be for this reason that recency has been neglected in theoretical treatments of learning. To evaluate the effects of recency, a situation is required in which one response will occur if frequency is the determining factor and another response will occur if recency is the important variable. A selective learning situation such as a simple T-maze should provide the necessary setting. Nevertheless, it should be noted that in many T-maze situations frequency and recency lead most often to identical predictions. Where reward is placed in one goal box on a high proportion of the trials, the animal soon responds predominantly to that side; consequently, the more frequent and most recent response are the same. A T-maze situation in which reward is randomly varied between the two alternatives at a 50-50 ratio should provide a



maximum of response variation. In such a situation the difference between frequency and recency should be most evident, and the predictive efficacy of the various theoretical models should be put to the most difficult test.

### EXPERIMENT I

This experiment was designed to compare the relative predictive efficacy of the theoretical models discussed above.

#### Procedure

Fifteen albino rats of the Sprague-Dawley strain were given 25 trials per day for two consecutive days on an elevated T maze. The noncorrection method was used. All animals were experimentally naive at the beginning of the experiment, and each was given pretraining on a straight runway. During the experimental trials, reward was randomly varied between the alternative goal boxes in a 50-50 ratio. The schedule of reward placement was constructed from a table of random numbers.

The maze was 4 ft. long by 2 in. wide with a starting arm 2 ft. by 2 in. The maze and two similar goal boxes were painted a uniform gray. In order to reduce distracting extra-maze cues, the experiment was conducted in the "dome" room of the Comparative Psychology Laboratory at the University of Texas. The experimental dome is 18 ft. in diameter with 5.5-ft. side walls arched to an 8-ft. center. It is constructed symmetrically of wooden ribs and covered with a fine mesh screen painted silver gray, hence giving a homogeneous surface radially symmetrical about the geometric center.

#### Analysis of Results

The problem was designed to determine which of the theoretical models discussed above most accurately predicts the individual responses of the fifteen Ss in this experiment. Since frequency and probability models require a minimal amount of past experience upon which prediction can be based, no predictions were made for the first five responses of each day. Thus, predictions were calculated for a total of 40 responses for each of the fifteen Ss. Table 1 compares the accuracy of predictions based upon each theoretical model.

The values entered in Column "*t*" of this table are *t*-test values for the differences between the mean accuracy of predictions based upon the specified model and the mean accuracy of predictions derived from the recency-weighted model. Since it was predicted that the latter theoretical model would prove to be a more accurate predictor of individual responses than any of the other models considered, a one-tailed test is appropriate. (See Jones, 8.) A *t* of 1.701 is significant at the .05 level, and a *t* of 2.467 is significant at the .01 level of confidence where 28 degrees of freedom are available. Predictions based upon the recency-weighted probability model were made by employ-

TABLE 1  
AVERAGE ACCURACY OF RESPONSE PREDICTIONS

Source of Prediction	Mean Accuracy	SD	<i>t</i>
	65.1	12.6	2.46
1. Postremity—"go where went last trial"	67.2	7.3	2.55
2. Postremity—"go where went last if rewarded, but alternate if not rewarded"	62.2	15.5	2.75
	64.8	13.8	2.06
3. Frequency of response			
4. Frequency of reward	61.2	10.7	3.75
5. Reaction potential—algebraic sum of positive and negative outcomes	67.2	11.3	2.11
6. Probability ( $p = \frac{\Sigma P}{\Sigma P + \Sigma N}$ )	75.2	9.7	—
7. Recency-weighted probability model			

ing two gross approximations. The only excuses for this procedure are found in the accuracy of predictions actually achieved, and in the likelihood that more precise coefficients should lead to even more accurate predictions. The first approximation involved the assignment of initial trace magnitudes to events in the learning sequence. The first rewarded or nonrewarded outcome in a succession of consecutive rewarded or nonrewarded outcomes was arbitrarily assigned a value of 100. Each successive reward following reward, or nonreward following nonreward, was arbitrarily assigned an initial trace magnitude which was 10 per cent less than the magnitude of the trace preceding it. When there was a shift from reward to nonreward or from nonreward to reward, the first outcome in the new sequence was assigned an initial magnitude of 100, and each successive outcome in this new run of reward or nonreward was reduced by 10 per cent. The second arbitrary approximation involved the selection of a 10 per cent decay coefficient. Upon the occurrence of each new outcome the magnitude of each residual trace was assumed to decay to 90 per cent of what it was before the occurrence of the new outcome. Through the accumulation of empirical evidence, it should be possible to correct and refine coefficients and to secure significantly more accurate predictions than those which resulted from this approximate method; nevertheless, a significant increase in accuracy has resulted from employing even this rough approximation of the theoretical model.

## EXPERIMENT II

A second experiment was designed to compare the relative efficacy of predictions based upon frequency and recency. Since no attempt was made to distinguish between postremity and the recency-weighted probability model tenta-

tively proposed here, the results of this experiment may also be interpreted as a test of the postremity principle in individual response prediction.

### *Procedure*

Thirty hooded rats were Ss in this experiment. Each was experimentally naive at the beginning of the experiment, and each was given pretraining on a straight runway. The apparatus was an elevated T maze 4 ft. in length with starting arm 2 ft. long. All units were 2 in. wide, and the maze surface was 15 in. from the floor. To eliminate extramaze cues, the experiment was conducted in the dome room described in Experiment I.

Each animal received all experimental trials on one day. Between each trial, Ss were returned to a restraining compartment for approximately 1 min. of rest. The noncorrection method was used throughout the experiment. On the experimental day, one trial was given to determine any initial preference; no reward was present in the maze during this trial. On succeeding training trials, reward was always present in one of the two goal boxes. A schedule of reward placement was constructed from a table of random numbers in such a manner that reward was predominantly on one side of the maze during the first half of training, but was then reversed so that it was presented predominantly on the other alternative for the latter half of training. Specifically, reward was presented in 9 out of the first 13 trials on the side initially preferred by the S, and then the schedule was reversed so that reward was presented in 9 out of the last 13 training trials on the side not initially preferred. Following the 26 trials in which reward was always present in one goal box, 15 additional trials were administered during which no reward was presented.

## RESULTS AND CONCLUSIONS

### *Criterion Measure*

An attempt was made to create a situation in which frequency and recency should lead to different predictions. Because reinforcement is postulated to have important effects upon the strengths of competing response tendencies, responses during extinction were selected as best for evaluating the

relative magnitudes of competing response tendencies.

The relative number of responses to each alternative during a complete "alternation cycle" has been utilized by Hull (7) as a measure of the relative strengths of computing reaction potentials. He states:

... The momentary oscillation principle does not prevent the occurrence of appreciable sequences of one reaction to the exclusion of the other. Perhaps the most fundamental concept ... is that of *response alternation*. A response alternation is said to occur when one type of response shifts to the other. For example, in the response-sequence fragment

$R- | R+ | R+ | R- | R- | R+$

there are three alternations, each marked by vertical lines. Our second concept, flowing directly from the first, is that of the *alternation phase*; this includes the number of reactions falling between two successive response alternations. Thus, in the above example, the first complete alternation phase represented contains three  $R+$ 's. Our third concept is that of the *alternation cycle*. An alternation cycle is the succession of responses comprised in two successive alternation phases. In the above example, an alternation cycle of  $3 + 2$ , or five reactions, is enclosed between the two heavy vertical lines. Finally, there is the concept of *asymmetry* of the response cycle; this refers to the fact that a behavior cycle may contain more reactions in one alternation phase than in the other (7, p. 42).

It is the asymmetry of responses during the first complete alternation cycle of the extinction period that is of especial interest to us here. Hull shows that where the reaction potential associated with one response is greater than that associated with the other, response alternation cycles will be asymmetrical, with a majority of responses made to the alternative associated with the greater reaction potential. Thus, the relative length of phases during an alternation cycle is an acceptable criterion for inferring relative strengths of competing response tendencies. Predictions based upon frequency and recency

principles were made for responses of the first complete alternation cycle of the extinction period, and a chi-square test used to determine the significance of number of correct predictions.

### Results

Twenty-one of the 30 Ss made a majority of responses, during the first complete alternation cycle of the extinction period, to the alternative on which *reward* had been most *recently* encountered. A chi square of 4.033 with 1 df is significant at the .05 level of confidence.

If it is predicted that a majority of responses during the first complete alternation cycle will be made to the alternative *responded* to most recently during training, it is found that the behavior of 22 of the 30 Ss is correctly predicted. With a one-tailed test, which is appropriate in light of the theoretical hypothesis of this paper, a chi square of 5.633 with 1 df is significant at the .01 level of confidence.

When it is predicted that a majority of responses during the first complete alternation cycle will be to the side *responded* to most *frequently*, the behavior of only 16 out of 30 Ss is correctly predicted.

The behavior of 17 out of 30 Ss is

TABLE 2  
RECENCY AND FREQUENCY PRINCIPLES IN  
PREDICTING MAJORITY OF RESPONSES  
DURING FIRST ALTERNATION CYCLE  
OF EXTINCTION PERIOD

Predict Majority of Responses	Correct	Incorrect
1. To alternative responded to most <i>recently</i>	22	8
2. To alternative rewarded most <i>recently</i>	21	9
3. To alternative responded to most <i>frequently</i>	16	14
4. To alternative rewarded most <i>frequently</i>	17	13



correctly predicted on the basis of *frequency of reward*. Because the total number of reward placements was equated for the two alternatives, the algebraic sum of positive and negative outcomes associated with responses to each side was always equal. Such a model would lead to the prediction of an equal probability for each alternative or, in other words, to no directional prediction in this situation.

### SUMMARY

1. One of the important controversies among contemporary learning theorists is the question of whether Ss utilize all past experiences equally or whether recent experiences are of greater importance in determining responses in a selective learning situation.

2. It is possible to utilize contemporary theories of learning as a basis for individual response predictions. Because learning is a change within the individual organism, the final test of a theory of learning is its ability to predict individual behavior.

3. The work of Voeks indicates that postremity predictions are less accurate where they conflict with frequency-based predictions than where the two predictions are not opposed, although postremity predictions were more accurate on the average in her situation. This suggests that a theoretical model which bases predictions upon *both* frequency and recency should be superior to one which utilizes only one principle or the other.

4. After reviewing models for individual response predictions which may be derived from postremity, frequency, and probability theories, a predictive model was tentatively suggested which incorporated both recency and frequency into a cognitive theory of learning.

5. Experiment I presented a T-maze situation in which reward was randomly varied (50-50) between the two alterna-

tives. Individual responses were predicted from postremity, frequency, and probability models. The accuracy of these predictions was compared to the accuracy of predictions based upon a model taking both frequency and recency into consideration. The latter model was found to predict with significantly higher accuracy than were any of the other models from which predictions were derived.

6. Experiment II presented a T-maze situation in which the reinforcement schedule first favored one alternative and was then reversed to favor the other alternative. The learning period was followed by an extinction period. The relative number of responses made to each alternative during the first complete alternation cycle of the extinction period was used as a measure of the relative strengths of competing response tendencies developed during the learning period. In this situation, it was found that recency predicted significantly better than chance, while frequency predicted only at the chance level.

7. The general conclusion of this paper is that a theoretical model which incorporates both the principles of recency and frequency can predict individual responses more accurately than one which considers only recency or frequency alone. Estes and Burke (5) have demonstrated that Guthrie's contiguous association theory can be quantified to serve as the basis for predictions which are based upon both recency and frequency. It should be possible for any of the contemporary theories of learning to incorporate both principles. Until frequency and probability theories include a recency principle it appears that they will be neglecting an important consideration, as shown by the growing body of data on the relative importance of recent events in the learning sequence.

8. This paper has indicated how a

cognitive-probability model might incorporate both recency and frequency principles. It is also suggested that refinement of predictions based upon this model can be achieved through developing coefficients which more accurately reflect empirical data.

## REFERENCES

1. BUSH, R. R., & MOSTELLER, F. A mathematical model for simple learning. *Psychol. Rev.*, 1951, 58, 313-323.
2. CRESPI, L. P. Quantitative variation of incentive and performance in the white rat. *Amer. J. Psychol.*, 1942, 55, 467-517.
3. CRESPI, L. P. Amount of reinforcement and level of performance. *Psychol. Rev.*, 1944, 51, 341-357.
4. ELLIOTT, M. H. The effect of change of reward on the maze performance of rats. *Univer. Calif. Publ. Psychol.*, 1928, 4, 19-30.
5. ESTES, W. K., & BURKE, C. J. Stimulus variability in learning. *Psychol. Rev.*, 1953, 60, 276-286.
6. GUTHRIE, E. R. *The psychology of learning*. New York: Harper, 1935.
7. HULL, C. L. *A behavior system: an introduction to behavior theory concerning the individual organism*. New Haven: Yale Univer. Press, 1952.
8. JONES, L. V. Test of hypotheses: one-sided vs. two-sided alternatives. *Psychol. Bull.*, 1952, 49, 43-46.
9. SPENCE, K. W. *Behavior theory and conditioning*. New Haven: Yale Univer. Press, 1956.
10. TINKLEPAUGH, O. L. An experimental study of representative factors in monkeys. *J. comp. Psychol.*, 1928, 8, 197-236.
11. TOLMAN, E. C. *Purposive behavior in animals and man*. New York: The Century Co., 1932.
12. TOLMAN, E. C., & BRUNSWIK, E. The organism and the causal texture of the environment. *Psychol. Rev.*, 1935, 42, 43-77.
13. VOEKS, V. W. Postremity, recency, and frequency as basis for prediction in the maze situation. *J. exp. Psychol.*, 1948, 38, 495-510.
14. VOEKS, V. W. A formalization and clarification of a theory of learning. *J. Psychol.*, 1950, 30, 341-362.

(Received March 30, 1957)

## AN ANALYSIS OF STIMULUS VARIABLES INFLUENCING THE 'PROPRIOCEPTIVE CONTROL OF MOVEMENTS

HARRY P. BAHRICK

*Ohio Wesleyan University*<sup>1, 2</sup>

It is generally known that accurate execution of movements depends upon proprioceptive information reaching the central nervous system. Clinical evidence (7, p. 235) as well as experimental findings (8) indicates that control and perception of movements are very poor when this sensory channel is not functioning.

Little is known, however, about the specific characteristics of proprioceptive stimulation that permit the individual to control changes in position, rate, or acceleration of his limbs. In other words, no detailed theories of proprioception comparable to the specific theories available for some other sense modalities have been developed (7, p. 234). Most of the available knowledge in this area is based upon anatomical investigations of the receptor system, its neural connections, and its central representations. Although several types of receptors have been identified (13, p. 1185; 14), differentiation of their function is as yet not clearly established. It is thought that forces internal to the body act as

proprioceptive stimuli, but the processes by which these stimuli are encoded into messages which ultimately form the basis for perception and control of movements are not well understood (7, p. 234).

Behavioral data specifying the relations between stimulus and response characteristics have been difficult to obtain because of problems of controlling proprioceptive stimuli. Investigators have used drugs or faradic currents (8, 9) as means of reducing the effectiveness of proprioceptive stimuli. Recently, an indirect approach to this problem has been attempted. This approach consists of varying the type and degree of resistance to motion offered by a control which *S* uses in the execution of movements. The effect of this variation upon *S*'s ability to perceive and control his movements is studied, and an attempt is made to infer characteristics of the proprioceptive system. As a technique of investigating proprioception, this approach has obvious limitations. The forces which *S* applies to move a control are only indirectly related to the proprioceptive stimulation he receives during the execution of the movement. The cutaneous senses are also stimulated during movement, and unknown transformations are involved between the control force acting upon the limb and the proximal stimuli acting upon receptors in muscles, tendons, or joints.

Despite these substantial limitations, the approach has some theoretical as well as practical advantages. The forces required to move a control, and thus also the control forces acting upon the

<sup>1</sup> This research was supported in part by the United States Air Force under Contract No. AF 41(657)-70 monitored by the Operator Laboratory, Air Force Personnel and Training Research Center, Randolph Air Force Base, Texas. Permission is granted for reproduction, translation, publication, use, and disposal in whole and in part by or for the United States Government.

<sup>2</sup> Work contributing to this theory was initiated in the Laboratory of Aviation Psychology of the Ohio State University in 1953, and is now being continued there. For the past year the author has continued research contributing to this theory at Ohio Wesleyan University through support from the National Science Foundation.



limb, can be specified as a function of four physical properties of the control, and these properties can all be regulated conveniently by  $E$ . They are mass, viscosity, elasticity, and the degree of coulomb friction. In a control such as that used by Howland and Noble (11) these parameters combine according to the following time-varying system equation:

$$L_t = K\theta + B d\theta/dt + J d^2\theta/dt^2, \quad (1)$$

where the left-hand side of the equation is the force applied by the human arm, the right-hand side represents the component resistive forces offered by the external control,  $L_t$  is the torque required to move the control at any instant of time ( $t$ ),  $K$  is the constant of elasticity of the control,  $B$  is the viscosity constant,  $J$  is the moment of inertia, and  $\theta$  is the angular displacement of the control with respect to its neutral, or spring-centered, position (6). Coulomb friction has been neglected in this equation, as have the internal resistive forces in the limb itself.

It has already been pointed out that the force which the control exerts upon the limb may not be equated to proprioceptive stimuli. However, one may assume that the forces which do act as proprioceptive stimuli during the movement of limbs are determined by physical properties of our limbs analogous to those specified in the above equation for the control. Previous investigation (4) has already established some of these physical properties of limbs and their significance in relation to the control of movements. These physical properties of limbs are difficult to control, and the present approach attempts to infer their function in proprioception by studying the effects of analogous characteristics of controls where these properties can be manipulated conveniently.

From an applied viewpoint, this approach may be useful in that the data are relevant to the solution of human

engineering problems related to the design of controls used in man-machine systems, or to the design of prosthetic devices.

In the present article some general hypotheses are developed about the effect of each of the physical control parameters specified in Equation 1, and data are reported which test these hypotheses.

Inspection of Equation 1 shows that the torque needed to move the control depends upon its position, rate, and angular acceleration, the relative importance of these depending upon the respective values of the elasticity, damping, and inertia constants. Thus, if the elasticity constant  $K$  is zero, the torque required to move the control will be independent of its position, but if  $K$  is relatively large, the torque will vary largely as a function of position. Analogous relations exist between the damping constant  $B$  and angular velocity, and between the moment of inertia and angular acceleration.

It is now hypothesized that a man can use the force cues obtained in moving the control to improve his perception of position, rate, and acceleration of limb motion. Specifically, it is hypothesized that the elasticity constant of the control improves  $S$ 's ability to perceive and control positions, the damping constant improves perception and control of rate, and the moment of inertia improves the perception and control of acceleration. Thus, an increase in each of these control constants should lead to improvement in the corresponding behavior. At the same time, it is hypothesized that an increase in any of the control constants will affect adversely performance which is aided by the other constants. Thus, increases in  $K$  are expected to interfere with the control and perception of rate and acceleration, while increases in  $B$  and  $J$  will affect adversely the control and perception of position. This hypothesis suggests itself, since the force

required to move the control would not be expected to provide useful cues for the control of rate if it changes rapidly with position, and, conversely, it should not offer useful cues for the detection of position if it varied greatly as a function of rate or acceleration.

Several experiments have been conducted in which the accuracy of movement was studied as a function of the physical characteristics of controls (1, 2, 3, 10, 11). A few of these (1, 2, 10) were designed specifically to test the above predictions. In one study (2) Ss performed simple circular and triangular control motions with a joystick control which was loaded with various degrees of spring stiffness, or damping, or mass. In each control-loading condition, the movements were first practiced with the help of a visual guide and paced by means of a metronome. The visual and auditory guides were then removed, and Ss were instructed to reproduce the motions as accurately as possible. Photographic records of all motions were obtained and measured for accuracy of temporal and spatial reproduction. It was found that an increase of viscous damping or of inertia of the control resulted in greater uniformity of speed within individual motions, and also in greater uniformity of speed in successive reproductions of the same motion. In the case of the triangular motions, increased mass and increased damping led to greater uniformity of peak velocity on each side of the triangle on successive trials. Spring loading interfered with the control of rate and acceleration, but its effect upon spatial accuracy of the reproduced motion was, in general, not significant. It was suggested that extended practice is needed for effective utilization of cues provided by spring loading.

This hypothesis was checked in a second experiment (1) in which the accu-

racy of positioning a horizontal arm control was investigated as a function of changes in the torque-displacement relation of the control. Extended practice was given and knowledge of results was provided. It was found that positioning errors are smallest when the ratio of relative torque change to displacement is largest. Under optimum conditions of spring loading, average positioning errors were less than half the amount obtained for a control which was not spring loaded. It was concluded that force cues provided by a spring-loaded control can improve the accuracy of positioning a control, and that the amount of improvement is a function of the relative and absolute torque change per unit of amplitude change.

Further investigation<sup>3</sup> of the usefulness of force cues in regulating the amplitude of motion has supported the above conclusions. It was shown that the transmission of amplitude information can be increased significantly by spring loading the control used by Ss. Optimum results were obtained with a control which provided geometric increments of force as a function of arithmetic changes of amplitude. This condition provides force cues which are equally discriminable over the range of amplitudes employed (12, 15), and yields the largest number of absolutely discriminable categories of amplitude response.

Although the above results support the general hypotheses regarding the effects of  $K$ ,  $B$ , and  $J$  constants upon the control of movements, many questions remain unanswered. In order to establish that the observed effects are due to changes in proprioceptive stimulation, it will be necessary to control cutaneous sensitivity. It is hypothesized here that

<sup>3</sup> Bahrack, H. P. Force cues and the control of movement amplitudes. (In preparation.)



the contribution of cutaneous receptors is most significant in relation to minute manipulatory responses, and least significant for larger movements of the type dealt with here.

Further problems arise because the control parameters under discussion have certain mechanical effects upon the nature of movements, and these must be separated from the effects upon proprioceptive stimulation. Large amounts of damping, for example, make rapid movements difficult and fatiguing, and greater uniformity of movement rate observed under these conditions may reflect mechanical effects rather than improved proprioceptive discrimination. The identification of these mechanical effects becomes more difficult when continuous movements are dealt with, as was shown in the study by Howland and Noble (11). Interactions among the physical parameters of controls may cause complex mechanical effects such as oscillation, and these may obscure or counteract the effects due to augmented proprioceptive stimulation. In general, the analysis of  $K$ ,  $B$ , and  $J$  effects is relatively simple for discrete, adjustive movements of the type primarily dealt with so far, but becomes increasingly complicated for complex or continuous motions.

Work now in progress attempts to establish relations between the forces exerted upon a control, and intensity of stimulation at the receptors in the elbow. This analysis is based upon a simplified mechanical model of the arm (16) by means of which forces acting upon the hand are resolved at the elbow joint (5, p. 319). In this manner it may become possible to infer changing intensities of stimulation of receptors at the joint during the course of movements.

Ultimately, the development of proprioceptive theory described here must be supported by a more direct analysis

of  $K$ ,  $B$ , and  $J$  factors within the body, and their effects upon the perception and control of movements. This, in turn, will require a better understanding of the biophysical principles by which forces internal to the body are brought to bear upon proprioceptive receptors.

# REFERENCES

1. BAHRICK, H. P., BENNETT, W. F., & FITTS, P. M. Accuracy of positioning responses as a function of spring loading in a control. *J. exp. Psychol.*, 1955, 49, 437-444.
2. BAHRICK, H. P., FITTS, P. M., & SCHNEIDER, R. The reproduction of simple movements as a function of proprioceptive feedback. *J. exp. Psychol.*, 1955, 49, 445-454.
3. DERWORT, A. Ueber die Formen unserer Bewegungen gegen verschiedenartige Widerstaende und ihre Bedeutung fuer die Wahrnehmung von Kraefften. *Z. f. Sinnesphysiol.*, 1943, 70, 135-183.
4. FENN, W. O. The mechanics of muscular contraction in man. *J. appl. Physics*, 1938, 9, 165-177.
5. FICK, R. *Handbuch der Anatomie und Mechanik der Gelenke*. Part II. Jena: Verlag von Gustav Fischer, 1910.
6. FITTS, P. M. Engineering psychology and equipment design. In S. S. Stevens (Ed.), *Handbook of experimental psychology*. New York: Wiley, 1951. Pp. 1287-1340.
7. GELDARD, F. A. *The human senses*. New York: Wiley, 1953.
8. GOLDSCHIEDER, A. Untersuchungen ueber den Muskelsinn. I. Ueber die Bewegungsempfindung. In A. Goldscheider, *Gesammelte Abhandlungen*, Vol. II. Leipzig: Barth, 1898.
9. GOLDSCHIEDER, A. Untersuchungen ueber den Muskelsinn. II. Ueber die Empfindung der Schwere und des Widerstandes. In A. Goldscheider, *Gesammelte Abhandlungen*, Vol. II. Leipzig: Barth, 1898.
10. HELSON, H., & HOWE, W. H. Inertia, friction, and diameter in handwheel tracking. OSRD Rep. No. 3454, 1943. (PB 406114.)
11. HOWLAND, D., & NOBLE, M. E. The effect of physical constants of a control



- on tracking performance. *J. exp. Psychol.*, 1953, 46, 353-360.
12. JENKINS, W. L. The discrimination and reproduction of motor adjustment with various types of aircraft controls. *Amer. J. Psychol.*, 1947, 60, 397-406.
13. JENKINS, W. L. Somesthesia. In S. S. Stevens (Ed.), *Handbook of experimental psychology*. New York: Wiley, 1951. Pp. 1172-1190.
14. MATHEWS, B. H. C. Nerve endings in mammalian muscle. *J. Physiol.*, 1933, 78, 1-53.
15. NOBLE, M. E., & BAHRICK, H. P. Response generalization as a function of intratask response similarity. *J. exp. Psychol.*, 1956, 51, 405-412.
16. WHITE, H. E. *Modern college physics*. New York: Van Nostrand, 1948.

(Received for early publication June 6, 1957)

# THE PSYCHOLOGICAL REVIEW

## UNCERTAINTY AND CONFLICT: A POINT OF CONTACT BETWEEN INFORMATION-THEORY AND BEHAVIOR-THEORY CONCEPTS<sup>1</sup>

D. E. BERLYNE

*Center for Advanced Study in the Behavioral Sciences<sup>2</sup>*

*Information theory*, originally designed to handle certain problems in communications engineering (41), needs to be distinguished from *psychological information theory*, which is one of its offshoots. The former consists of a mathematical language, incorporating a number of distinctive measuring techniques. Psychological information theory is, in contrast, a type of theory in the scientific sense: it applies information-theory measures to phenomena within the purview of psychology and uses information-theory language to formulate laws or hypotheses with testable implications about behavior.

Recent literature contains several sketches of such theory (e.g., 1, 26, 35), mostly concerned with how human beings code information or with how much information can pass through them in particular situations. There are many unmistakable affinities between this kind of psychological theory and S-R behavior theory (learning theory): they have overlapping interests in

such matters as discrimination, remembering and reaction time, they share a predilection for operationally definable and quantitative concepts, and they start out, respectively, from the closely related "black-box" and "neobehaviorist" points of view. It is therefore rather disappointing that so little integration between theories of the two types has yet taken place. We can regard two theories as "integrated" if one can be deduced from the other or if both can be deduced from a third theory. Before any integration can be attempted, the scope of information-theory language within the domain of behavior theory must be examined, which means considering to what extent recourse to it is *possible* and to what extent, if possible, it is *useful*.

The use of information-theory measures is *possible* whenever we have a *partition*, i.e., a set of phenomena that can be divided into non-overlapping subsets, and a *probability distribution*, i.e., a way of associating with each subset a number from 0 to 1, such that the numbers associated with all the subsets in the partition add up to 1. Whenever these two requirements are fulfilled, such measures as "amount of information," "uncertainty," and "relative uncertainty" can be applied. As

<sup>1</sup> This article owes a great deal to discussions with Dr. I. R. Savage and several other colleagues at the Center for Advanced Study in the Behavioral Sciences. It was written while the author was on leave of absence from the University of Aberdeen, Scotland.

<sup>2</sup> Now at the University of California, Berkeley.

soon as we have *two* sets of phenomena satisfying both requirements, the two can jointly be regarded as a "transducer," and the relations between them described in terms of "transmitted information," "noise," and "equivocation."

The phenomena that concern behavior theory consist, in fact, of two sets that can be partitioned into subsets with associated probabilities, namely *stimuli* and *responses*. The language of information theory is therefore, in principle, applicable to everything within the competence of behavior theory. Two limitations to its appropriateness have often been pointed out (14, 47). One is that the behavior theorist is especially interested in learning, i.e., in situations where probabilities of responses are changing. The other is that information-theory measures take no account of any ordering of the subsets in the partition or, more particularly, of the fact that stimuli and responses are not confined to nominal scaling (44). But these limitations are not insuperable. Information-theory measures can be derived from response probabilities at different stages of a learning process and compared, or else they can be applied when learning is near its asymptote. Stimuli and responses can be successively described in information-theory terms and in terms of physical or psychophysical dimensions, and the different measures can be related. For example, a response-class can have attached to it both a mean reaction time and a rate of transmitted information, and connections between the two can be explored.

If the possibility of describing the domain of behavior theory in information-theory language is accepted, the question of its *usefulness* still remains. One of the principal functions of any language is to make secondary or mediated generalization and discrimination possible. A language incorporates classifications, of which measures are spe-

cial cases. Classifications are procedures for attaching certain descriptive terms (values in the case of measures) as verbal responses to certain items in the universe of discourse but not to others. Items bearing a common verbal label come to evoke similar behavior in the users of the language. A classification is useful only as long as the items allotted the same label share some important quality, such that a common response to them will be rewarded (or reinforced) despite other qualities that might distinguish them. Information-theory measures are useful for the description of behavior, therefore, if these measures are closely related to other variables that have proved to be important for psychology.

A large body of data demonstrating that such is the case has been amassed within the last ten years. Reaction time, retention of verbal material, and accuracy of psychophysical judgment, to cite examples, appear to be functions of "uncertainty" and "amount of transmitted information." The situations in which such associations have been found have, however, been situations in which *subjects have some knowledge of the range of alternative stimuli that might occur and of their probabilities*. This knowledge is provided by *E's* instructions, or by the presentation of a sufficient sample of material for estimates to be made, or, as in experiments using natural languages, by previous training. It has, indeed, been contended by Cronbach (14) that information-theory measures in psychology should be confined to cases where "the receiver knows the probabilities and joint probabilities of the source." As Cherry reminds us, information theory is part of the "metalanguage of an external observer; it is not a description of the process of communication as it appears to one of the participants" (13, p. 170). An observer can compute information-theory meas-



ures from data not accessible to the individuals he is observing. But there is not likely to be much connection between these measures and variables of psychological importance, unless there is some isomorphism between the situation as viewed by the observer and the situation as it impinges on the observed organism.

The situations in which the use of information-theory terminology has had some success can be analyzed further as follows:

1. There is an antecedent stimulus-pattern,  $S_x$ . It may consist of the background conditions of the experiment, of an  $E$ 's warning signal or, in sequential studies, of any item in a sequence.

2. Whenever  $S_x$  occurs, it is followed by one and only one of a set of consequent stimuli  $\{S_1 \dots S_n\}$ .

3. Whenever one of the consequent stimuli occurs, a particular response corresponding to it is performed.

4. The responses corresponding to the consequent stimuli are such that no more than one of them can be performed at once, whether because of the  $E$ 's instructions or because of some physiological incompatibility between them.

In such situations, one can predict that all the  $n$  responses corresponding to the  $n$  consequent stimuli will become conditioned to  $S_x$ . No more than fractional components of these responses can be expected to occur immediately after the onset of  $S_x$ , both because simultaneous performance of the complete responses is precluded by the conditions of the experiment and because performance of any of them before the consequent stimulus appears will not be reinforced, so that the conditions for inhibition of delay will be fulfilled (38).  $S_x$  will thus come to evoke *competing response tendencies*. For Hull's theory

(27, 28), these response tendencies will be "reaction potentials." Cognitive behavior theories (e.g., 45) would describe them as "expectations" of the consequent stimuli, and the "expectation" resembles the "reaction potential" insofar as both imply the occurrence of a particular response, if certain additional conditions are met.

Furthermore, the relative *strengths* of the competing response tendencies will reflect the probabilities of the corresponding stimuli. Whether one regards the number of reinforced trials (27, 28), the variety of stimulus situations that have been contiguous with the response (23), or the number of times an expectation has been confirmed (45) as the decisive factor, responses associated with more frequent consequent stimuli will become more strongly associated with  $S_x$ . There is, in fact, experimental evidence (17, 21) that the strength of a predictive verbal response (which is especially relevant here), as judged by the asymptote of response probability, increases with the probability of the corresponding stimulus.

To sum up, the situations in which information-theory language has been of value are ones in which *conflict* is an important factor, and the theory of conflict seems to be one area where linkages between information theory and behavior theory may hopefully be sought.

#### DEGREE OF CONFLICT (C)

If the study of conflict is to progress beyond noting the effects of its presence or absence, some way of distinguishing *degrees of conflict* will have to be adopted. The degree of conflict is, of course, not necessarily the same as the *severity of the effects of conflict*, of which it is likely to be merely one determinant. Other determinants would be the nature of the conflicting response tendencies (e.g., whether they are approach or avoidance tendencies [36])

and the conflict tolerance of an individual organism. Brown and Farber (11) suggest two conditions for the degree-of-conflict function (or, as they call it, "frustration"), viz., that it increase with the *absolute strengths* of the competing tendencies, and that it increase as their strengths approach *equality*. They, like most writers who have considered psychological conflict, confine their attention to conflicts between two response tendencies. If conflicts involving three or more alternatives are to be included in the treatment, as would seem desirable, the *number* of competing tendencies can be proposed as a third variable with which degree of conflict increases (4).

We can thus state the conditions for a degree-of-conflict function a little more precisely, as follows. Let us assume that response tendencies corresponding to a set of responses  $\{R_1 \dots R_n\}$  occur in an organism, that the responses in the set are such that no two of them can be performed at once, and that some non-negative quantity  $E$  (e.g., Hull's "reaction potential") can be associated with each response tendency (as a measure of its strength).

It should be noted that, although the *responses* cannot occur simultaneously, we are assuming that their corresponding *response tendencies* can. Second, there is no reason why several independent sets of competing response tendencies should not be aroused in the same organism at once. Third, we are considering cases where there is complete incompatibility, whether innate or learned, between alternative responses. It is, however, conceivable that two responses may be *partially* antagonistic, i.e., the evocation of one may reduce the amplitude or probability of the other without excluding its performance altogether. This may suggest *degree of incompatibility* between responses as an additional determinant of degree of conflict (4), which would complicate any mathematical treatment. Possible ways

of reducing degree of incompatibility to other variables, when learned incompatibility is involved, are considered elsewhere (7).

The degree-of-conflict function  $C(E_1 \dots E_n)$  should then have the following properties:

1.  $C$  is continuous and symmetric in the  $E_i$ ;
2.  $C \geq 0$ ;
3. if  $n = 1$ ,  $C = 0$ ;
4. with  $\sum_{i=1}^n E_i$  held constant,  $C$  reaches an absolute maximum when  $E_1 = E_2 = \dots = E_n$ ;
5. if  $E_1 = E_2 = \dots = E_n$ , and a response  $R_{n+1}$  with strength  $E_{n+1} > 0$  is added to the set  $\{R_1 \dots R_n\}$ ,  $C$  increases;
6. if every  $E_i$  is multiplied by  $k > 1$ ,  $C$  increases.

Now, let us suppose that we have a way of translating the  $E$  value for each response into a  $p$  value, or measure of probability, such that

1.  $0 \leq p \leq 1$ ;
2.  $\sum_{i=1}^n p_i = 1$ ;
3. if  $E_1 = E_2 = \dots = E_n$ , then  $p_1 = p_2 = \dots = p_n = 1/n$ ;
4. if one  $E_i$  increases with the others held constant, then the corresponding  $p$  increases and the other  $p$ 's decrease.

Some theorists (e.g., 12, 16) content themselves with probability as a sole measure of response strength. Others (e.g., 27, 42, 43) recognize additional ones, such as latency, frequency, resistance to extinction, amplitude and vigor. Of these, mean latency and mean frequency are merely the reciprocal of the probability that a response of the class in question will occur during one unit of time. Resistance to extinction may be regarded as the rate at which response



probability decreases when reinforcement is withdrawn. But other measures of response strength, expressing the energy with which the response is performed, are not the same as probabilities. Hull (28, pp. 25 ff.) and Spence (43, App. A) present methods for transforming  $E$ s into probabilities, when  $E$  is the Hullian reaction potential. Doing this means, however, losing information, since many sets of  $E$  values can be represented by the same set of  $p$  values. Whenever we have two or more independently defined response classes, as distinct from one response class and its complement, probability is a measure of *relative* and not absolute response strength. The distinction may be important. For example, Mr. A. may be torn between his duty to the community and his duty to his family, while Mr. B may have difficulty in deciding whether or not to spend a small sum on a newspaper. Both of them have two response tendencies with probabilities of .5, but in other respects the effects of the two conflicts may be radically different.

Be that as it may, the use of probabilities to express response strengths provides us with the partition and the probability distribution that are necessary conditions for recourse to information-theory measures. And if we examine the information theorist's formula for "uncertainty" or "entropy" ( $-\sum_i p_i \log_2 p_i$ ), we find that it satisfies the first five of our requirements for a degree-of-conflict function, but not the sixth. It increases with the number of alternative responses and is at a maximum when their strengths are equal. But it does not vary with their absolute strengths. In order to make "uncertainty" fulfill all our conditions, we can multiply it by some such quantity as the mean  $E$ . French's hypothesis (20) that the frustrating effects of a binary conflict are a function of the weaker of the two opposing forces suggests that "uncertainty" should be multiplied by the minimum rather than the mean  $E$ . But this would produce rather anomalous

results in higher-order conflicts when there are one very weak and several very strong response tendencies in competition. Our expression for degree of conflict then becomes  $-\bar{E} \sum_i p_i \log p_i \cdots 1$ . Put somewhat differently, "uncertainty" can be regarded as an indication of the "complexity" of a conflict, or of the difficulty that an observer would have in predicting which of the conflicting responses will be the first to occur. It does not reflect the "scale" of the conflict, which depends on the energy invested in the competing response tendencies. There may be a temptation to relate these two components to the *utility* and *probability-of-outcome* factors that must be taken into account in decision theory, or to the *motivational* and *structural* factors that often have been distinguished in psychological literature. But any such correspondence would be misleading. Both the "uncertainty" and the  $\bar{E}$  are determined by absolute response strengths, which depend on both motivational (utility) and structural (probability-of-outcome) variables; e.g., Hull's "reaction potential" (28) depends on "drive" and "amount of reinforcement" on the one hand and on "number of reinforcements" (habit-strength) on the other. It is interesting to observe that Shannon (41, p. 19) gives  $-K \sum_i p_i \log p_i$  as the only function satisfying his assumptions, and goes on to describe  $K$  as amounting to a "choice of a unit of measure" or, in other words, to some scaling factor comparable to our  $\bar{E}$ .

Expression 1 is, however, by no means the only one that will accord with our requirements. Another function, for example, that will do so without necessitating a transformation of  $E$ , is  $\sum_i (\log (\sum_i E_i) - \log E_i) \cdots 2$ . If  $E$  represents Hull's reaction potential, this function will not, in general, have the same values as Expression 1, because probabilities are not proportional to reaction potentials. It will, however, be an increasing monotonic transform of Expression 1.



We are not even confined to logarithmic functions, since we lack the additivity requirement that makes them mandatory for Shannon's purposes. A non-logarithmic function that will pass muster is

$$\frac{(\sum E_i)^2(n-1)}{1 + \sum_i (E_i - \bar{E})^2} \cdots 3.$$

Our requirements are, in fact, very weak ones, which a large number of functions will fit. Further research will, no doubt, add stipulations, allowing the range of possible functions to be narrowed down. For instance, one additional requirement that may be held reasonable, in view both of everyday observation of persons confronted with choices and of the logarithmic relation that obtains between number of alternative stimuli and choice reaction time, is that  $C$  should be a negatively accelerated increasing function of  $n$ . If this were adopted, then Expression 1 would be among those still meriting consideration, but Expressions 2 and 3 would be ruled out.

#### CORRELATES OF DEGREE OF CONFLICT

A degree-of-conflict measure, like an information-theory measure, can be justified as a classificatory device only if situations that have a common value assigned to them by the measure result in similar behavior, much as they may differ in other respects. The following are some psychological variables that appear likely, in the light of present knowledge, to depend on degree of conflict. They may actually turn out to be closely interrelated, but they are here separated for convenience.

1. *Emotional disturbance.* Various writers, from Dewey (15) on, have mentioned conflict as a cause of "emotion." Both the special reaction patterns (24) and the disruption of habitual behavior (31) that are characteristic of "emotional disturbance" have been ascribed to the occurrence of divergent neural processes. The power of conflict to precipitate neurotic behavior was pointed out independently by Pavlov and by

Freud, employing very different research techniques. So far, merely the dependence of these phenomena on relatively severe conflict has been noted, but future progress may well demand a quantitative treatment, in which intensities of disturbance are differentiated and related to degrees of conflict.

2. *Reaction time.* A lengthening of reaction time (or decision time or choice time) has often been reported as a consequence of conflict (see Berlyne [7]). A link with information theory presents itself in the finding that reaction time increases linearly, at least in some conditions, with "uncertainty" (26, 29); mean reaction time has been found to increase when alternative stimuli approach equiprobability and when they become more numerous. If, as we concluded, the number of competing response tendencies corresponds to the number of alternative stimuli, and if the relative strengths of those tendencies reflect the probabilities of the corresponding stimuli, we can infer that two of the suggested determinants of  $C$  affect reaction time.

Both traditional experimental psychology and psychological information theory have hitherto concentrated on "*forced-choice*" situations, in which only one response is appropriate to each alternative stimulus, and selection of a response depends on identification of the stimulus. A recent investigation by the writer (7) compared forced choices with *free choices*. For the latter, two or more stimuli were presented together, and the response corresponding to any one of them was to be performed. Both kinds of choice can be assumed to entail conflict: the free choice means a conflict between response tendencies of about equal strength evoked by the stimuli that are simultaneously present, while the forced choice means an unequal and therefore relatively mild conflict between a strong tendency to respond correctly to the one stimulus that occurs and weak tendencies to make re-

sponses appropriate to other stimuli, resulting from generalization. The usual information-theory analysis of the forced choice, in which the *S* is viewed as a transducer with a limited channel capacity, is not helpful for the treatment of the free choice.

Free-choice reaction times invariably exceeded forced-choice reaction times, and both were longer when the number of alternative stimuli and responses was increased from two to four, as the hypothesis that reaction time increases with degree of conflict would lead one to expect. Furthermore, when the absolute strengths of the response tendencies—the determinant of degree of conflict that is disregarded by uncertainty—were manipulated by changing the intensity or extensity of the stimuli, changes in free-choice reaction time resulted.

3. *Drive*. Various considerations and observed phenomena have led a number of writers (e.g., 11, 30, 46) to conclude that conflict may be a drive condition. The drive resulting from conflict as such must, of course, be distinguished from other drives that may be at a high level because conflict blocks the behavior that would normally reduce them.

A certain amount of evidence for a conflict drive was obtained by Lowell (30), who found approach-approach conflict to produce a greater speed of running in rats than a single approach tendency. A supplementary observation fitting our conception of *C* was that the conflict drive was not so much in evidence when the stimuli were unequal in intensity or when learning was incomplete (and the competing response tendencies presumably relatively weak).

Wyckoff's experiment (48, 49) provides other data that might be predicted from our assumptions. His pigeon *Ss* were rewarded with food when they pecked at a key of a certain color and not rewarded when the key was of another color. They were then tested with the key white, but the color indi-

cating whether pecking would be reinforced or not appeared if the animal stepped on a pedal. The pedal response was rapidly learned, even though it did not affect the probability of receiving food. It merely diminished the pigeon's "uncertainty" by one bit. The white key is reminiscent of the stimulus that made the dog neurotic in the famous Shenger-Krestovnikova experiment (38, pp. 290 ff.). This stimulus, intermediate in shape between the reinforced circle and the nonreinforced ellipse, was thought by Pavlov (38, p. 318), to produce a "conflict between excitation and inhibition." If the white key produced a conflict in Wyckoff's pigeons between tendencies to peck and to refrain from pecking, or between tendencies to expect and not to expect food, the coloring of the key that was a consequence of stepping on the pedal must have reduced the conflict by strengthening one response tendency and weakening the other. If a conflict drive is proportional to *C*, reduction of the drive can be expected to reinforce the pedal-stepping response. When the discrimination was reversed, Wyckoff found that the frequency of the pedal response would temporarily decrease. This also fits our interpretation, as each color would then go through a stage of evoking both tendencies, and seeing the colored key would thus increase rather than reduce conflict. Wyckoff himself offers an alternative explanation in terms of secondary reinforcement, but this leads into difficulties, as Prokasy points out in his report of a somewhat similar experiment (39).

Yet another relevant experiment is one by Fonberg (19). She trained dogs to perform a certain response ( $R_1$ ) as a way of terminating stimuli that had been associated with puffs of air or electric shocks. The animals then received training in quite a different response ( $R_2$ ), which was followed by food reinforcement in the presence of a loud tone but not in the presence of a



faint tone. When they were later subjected to a Shenger-Krestovnikova type of conflict by exposure to tones intermediate in intensity between the positive and negative alimentary conditioned stimuli, they reverted to their defensive response ( $R_1$ ). This finding indicates that the physiological state produced by a conflict, even when noxious stimuli have played no part in it, may be sufficiently similar to the physiological state (fear or anxiety) resulting from a noxious stimulus for generalization between the two to occur.

4. *Curiosity*. There is currently a good deal of interest in certain sorts of behavior whose main function seems to be the provision of information, and information theory might reasonably be expected to throw some light on them. The behavior under discussion includes the "exploratory" activities that bring about opportunities to perceive objects more readily; the verbal activities, including asking questions, that elicit informative verbal behavior from other individuals; and the symbolic activities that allow thought processes to feed on information other than that supplied by the immediate environment.

"Novelty" has often been mentioned as a distinguishing mark of situations that provoke such activities (2, 6). But something can either be relatively novel, in the sense that it has never been encountered before in its present context, or absolutely novel, in the sense that it has never been encountered at all. In both cases, we have situations in which "amount of information" is high, since this measure is inversely related to the probability of an event, and the probability of particular novel occurrence must be low in the light of an individual's past experience. We can also speak of conflict in connection with the same occurrences. A relatively novel stimulus pattern is one in which perception conflicts with the expectations aroused by the context. Moreover, at least as far as human beings are con-

cerned, any absolutely novel object is bound to consist of an unfamiliar combination of familiar elements or to possess characteristics intermediate between those of several well-known objects. Such an object can be expected to induce conflict, since it will inevitably evoke, by generalization, responses appropriate to a number of discrepant familiar objects.

Other words that seem apposite to situations that call for investigatory behavior are "doubt," "perplexity," and "ambiguity." These words likewise imply some degree of behavioral conflict; they indicate that different aspects of a situation evoke discordant reactions or else that a particular reaction is called forth by one aspect and inhibited by another. They are opposite in meaning to words like "clear" and "distinct," which generally imply that certain response tendencies have come, through discriminatory learning, to predominate over their competitors. "Doubtful," "perplexing," or "ambiguous" stimulus situations are usually also cases of high "uncertainty" in the information-theory sense, both because the subject cannot predict very successfully what the future behavior or the hidden properties of the entities will be, and because observers will not be able to predict very successfully how he will react to them. Nevertheless, curiosity is by no means always commensurate with "uncertainty"; there are many events whose outcomes are uncertain and yet which leave us completely indifferent. For knowledge of the outcome to be rewarding, the event must be of some "interest" to us, which usually means that strong habits or drives must be aroused. In other words, curiosity seems to be a matter of conflict rather than of "uncertainty" alone; "uncertainty" may be high, but there will not be much conflict if the absolute response strengths are low. That human beings, like Wyckoff's pigeons, find relief from doubt about vital matters rewarding, even when the



truth is unpleasant, is attested by common experience. Of the convicts studied by Farber (18), those who did not know how much time they would have to serve suffered more than those who were certain that they would never be paroled.

The writer suggested a few years ago (4) that at least some forms of human curiosity spring from the drive-producing properties of conflict. The conflicts that seem especially pertinent are those between implicit, most often symbolic, responses, such as "beliefs," "ways of thinking," and "ways of perceiving," whose incompatibility is largely an effect of learning. There are experimental data supporting the conclusion that curiosity, measured in various ways, is an increasing function of  $C$  (5, 8, 9).

5. *Stimulus complexity.* Among the various properties by which stimulus patterns can be classified, there is a group that can only be described collectively by some such term as "complexity." They are hard to define rigorously, and a number of quite distinct dimensions will, in all likelihood, be unraveled by attempts to do so. But the influence of this aspect of perceived material is revealed in several contexts: the special properties attributed to less complex (more "pragmatic") figures by the Gestalt school, the bearing of degree of complexity on aesthetic preferences, and, more recently, the influence of stimulus complexity on exploratory behavior on animals (see 10).

Attneave (1) has related the "complexity" dimension in visual figures to information theory through the concept of "redundancy," the inverse of "relative uncertainty." His treatment suggests a possible link between these variables and conflict. More "complex" stimulus patterns might well be those arousing more conflict, e.g., between perception of one part and expectations or redintegrative perceptual responses (3, 24, 37) aroused by other parts, between verbal or other classificatory re-

sponses, or between ocular and other orienting movements. If this hypothesis is well founded, we should expect more "complex" (or less "redundant") figures, like figures arousing conflict in other ways, to elicit more investigatory behavior. Experimental data confirming this prediction are available (8, 9).

6. *Reward.* While the punishing or drive-producing role of conflict is more evident and has received more attention, the possibility that conflict and uncertainty may at times be rewarding is suggested by gambling and aesthetic behavior. Similarly, journalistic practice seems to indicate a positive relation between the reward value of a piece of news and the "amount of information" it contains, which depends on its improbability or surprisingness (40). Surprise, like novelty, seems to mean some sort of clash between the reactions occasioned by an unexpected situation and those evoked anticipatorily through previously established habits (3). Surprising statements are, at least in certain circumstances, recalled more readily than others (5), and maze-learning experiments (see 10) show that exposure to a more complex environment (which, as we have seen, may mean a more conflictful environment) can be more reinforcing than exposure to a simpler one.

If conflict is usually an aversive condition but occasionally functions as a reward, it resembles fear, which likewise seems to be actively sought at times, e.g., at fairgrounds and in dangerous sports. The analogy with fear suggests two hypotheses to account for the paradox. One is that drive arousal may be rewarding at a moderate level. Hebb refers to "the *positive attraction of risk taking*, or mild fear, and of *problem solving*, or mild frustration," and speculates that "when arousal or drive is at a low level, . . . a response that produces increased stimulation and greater arousal will tend to be repeated" (25, p. 250). McClelland *et al.* (33) propound a rather similar hypothesis, whose bear-

ing on conflict is a little more conspicuous: "positive affect is the result of smaller discrepancies of a sensory or perceptual event from the adaptation of the organism; negative affect is the result of larger discrepancies." There have been a number of recent studies (e.g., Marx *et al.* [32]) showing that an increase in illumination up to a certain intensity will reinforce a bar-pressing response in a rat, while light of much greater intensity is known to be aversive.

The second hypothesis is that such states as fear or conflict are sought only when their arousal in similar circumstances has reliably and speedily been followed by drive reduction in the past.

Two recent works by empirically minded aestheticians provide some corroboration for these hypotheses. Graves (22) contends that the appeal of a visual design depends on variety, but that one part or quality must be made to dominate the others if the effect is to be satisfying. This would keep within bounds any conflict aroused. In accord with the second hypothesis, Meyer (34) shows that music owes much of its savor to continual departures from what preceding or accompanying patterns lead the listener to expect. But what is initially heard as an incongruity is invested with a new meaning by what follows, so that the momentary conflict is promptly resolved.

### SUMMARY

The use of information-theory measures is possible whenever there is a partition and a probability distribution. The stimuli and responses of behavior theory fulfill these conditions, but the situations in which information-theory language has proved useful to psychology have been ones in which conflict is an important factor. The "uncertainty" function satisfies some of the requirements that may reasonably be laid down for a measure of "degree of conflict." But it does not satisfy them all with-

out some modification, because it depends on the relative but not the absolute strengths of competing response tendencies.

A discussion of six psychological variables that appear to depend on degree of conflict reveals several further links with information theory. The variables are emotional disturbance, reaction time, drive, curiosity, stimulus complexity, and reward.

### REFERENCES

1. ATTNEAVE, F. Some informational aspects of visual perception. *Psychol. Rev.*, 1954, 61, 183-193.
2. BERLYNE, D. E. Novelty and curiosity as determinants of exploratory behavior. *Brit. J. Psychol.*, 1950, 41, 68-80.
3. BERLYNE, D. E. Knowledge and stimulus-response psychology. *Psychol. Rev.*, 1954, 61, 245-254.
4. BERLYNE, D. E. A theory of human curiosity. *Brit. J. Psychol.*, 1954, 45, 180-191.
5. BERLYNE, D. E. An experimental study of human curiosity. *Brit. J. Psychol.*, 1954, 45, 256-265.
6. BERLYNE, D. E. The arousal and satiation of perceptual curiosity in the rat. *J. comp. physiol. Psychol.*, 1955, 48, 238-246.
7. BERLYNE, D. E. Conflict and choice time. *Brit. J. Psychol.*, 1957, 48, 106-118.
8. BERLYNE, D. E. Conflict and information-theory variables as determinants of human perceptual curiosity. *J. exp. Psychol.*, 1957, 53, 399-404.
9. BERLYNE, D. E. The influence of complexity and change in visual figures on orienting responses. *J. exp. Psychol.*, 1958 (in press).
10. BERLYNE, D. E., & SLATER, J. Perceptual curiosity, exploratory behavior and maze learning. *J. comp. physiol. Psychol.*, 1957, 50, 228-232.
11. BROWN, J. S., & FARBER, I. E. Emotions conceptualized as intervening variables—with suggestions toward a theory of frustration. *Psychol. Bull.*, 1951, 48, 465-495.
12. BUSH, R. R., & MOSTELLER, F. *Stochastic models for learning*. New York: Wiley, 1955.
13. CHERRY, C. *On human communication*. New York: Wiley, 1957.
14. CRONBACH, L. J. On the non-rational application of information measures in psychology. In H. Quastler (Ed.), *In-*



- formation theory in psychology. Glencoe, Ill.: Free Press, 1955.
15. DEWEY, J. The theory of emotion. II. The significance of emotion. *Psychol. Rev.*, 1895, 2, 13-32.
  16. ESTES, W. K. Toward a statistical theory of learning. *Psychol. Rev.*, 1950, 57, 94-107.
  17. ESTES, W. K. Individual behavior in uncertain situations: an interpretation in terms of statistical association theory. In R. M. Thrall *et al.* (Eds.), *Decision processes*. New York: Wiley, 1954.
  18. FARBER, M. L. Suffering and the time perspective of the prisoner. *Univ. of Iowa Studies in Child Welfare*, 1944, 20, 155-227.
  19. FONBERG, E. On the manifestation of conditioned defensive reactions in stress. *Bull. Soc. Sci. et Lett. de Eds.*, 1956, 7, 1-8.
  20. FRENCH, J. R. P. Organized and unorganized groups under fear and frustration. *Univ. of Iowa Studies in Child Welfare*, 1944, 20, 231-308.
  21. GRANT, D. A., HAKE, H. W., & HORNSETH, P. Acquisition and extinction of a verbal conditioned response with differing percentages of reinforcement. *J. exp. Psychol.*, 1951, 42, 1-5.
  22. GRAVES, M. E. *Art of color and design*. (2nd ed.) New York: McGraw-Hill, 1951.
  23. GUTHRIE, E. R. *The psychology of learning*. New York: Harper, 1935.
  24. HEBB, D. O. *The organization of behavior*. New York: Wiley, 1949.
  25. HEBB, D. O. Drives and the C.N.S. (conceptual nervous system). *Psychol. Rev.*, 1955, 62, 243-254.
  26. HICKS, W. E. On the rate of gain of information. *Quart. J. exp. Psychol.*, 1952, 4, 11-26.
  27. HULL, C. L. *Principles of behavior*. New York: Appleton-Century, 1943.
  28. HULL, C. L. *A behavior system*. New Haven: Yale Univ. Press, 1952.
  29. HYMAN, R. Stimulus information as a determinant of reaction time. *J. exp. Psychol.*, 1953, 45, 188-196.
  30. LOWELL, E. L. The effect of conflict on motivation. Unpublished Ph.D. thesis, Harvard Univ., 1952.
  31. LURIA, A. R. *The nature of human conflicts*. New York: Liveright, 1932.
  32. MARK, M. H., HENDERSON, R. L., & ROBERTS, C. L. Positive reinforcement of the bar-pressing response by a light stimulus following dark operant pretests with no aftereffect. *J. comp. physiol. Psychol.*, 1955, 48, 73-76.
  33. MCCLELLAND, D., ATKINSON, J. W., CLARK, R. A., & LOWELL, E. L. *The achievement motive*. New York: Appleton-Century-Crofts, 1953.
  34. MEYER, L. B. *Emotion and meaning in music*. Chicago: Univ. of Chicago Press, 1956.
  35. MILLER, G. A. The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychol. Rev.*, 1956, 63, 81-97.
  36. MILLER, N. E. Experimental studies of conflict. In J. McV. Hunt (Ed.), *Personality and the behavior disorders*. New York: Ronald, 1944.
  37. OSGOOD, C. E. Behavior theory and the social sciences. *Behav. Sci.*, 1956, 1, 167-185.
  38. PAVLOV, I. P. *Conditioned Reflexes*. Oxford: Oxford Univ. Press, 1927.
  39. PROKASY, W. F. The acquisition of observing responses in the absence of differential external reinforcement. *J. comp. physiol. Psychol.*, 1956, 49, 131-134.
  40. SAMSON, E. W. *Fundamental natural concepts of information theory*. AFCL Rep. E5079, 1951.
  41. SHANNON, C. E., & WEAVER, W. *The mathematical theory of communication*. Urbana: Univ. of Illinois Press, 1949.
  42. SKINNER, B. F. *The behavior of organisms*. New York: Appleton-Century, 1938.
  43. SPENCE, K. W. *Behavior theory and conditioning*. New Haven: Yale Univ. Press, 1956.
  44. STEVENS, S. S. Mathematics, measurement and psychophysics. In S. S. Stevens (Ed.), *Handbook of experimental psychology*. New York: Wiley, 1951.
  45. TOLMAN, E. C. *Purposive behavior in animals and men*. New York: Appleton-Century, 1932.
  46. WHITING, J. W. M., & CHILD, I. L. *Child training and personality*. New Haven: Yale Univ. Press, 1953.
  47. WILSON, K. The information theory approach. In C. E. Osgood & T. A. Sebeok (Eds.), *Psycholinguistics*. *J. abnorm. soc. Psychol.*, 1954, 49 (4, Pt. 2—Suppl.), 35-49.
  48. WYCKOFF, L. B. The role of observing responses in discrimination learning: Part I. *Psychol. Rev.*, 1952, 59, 431-442.
  49. WYCKOFF, L. B. The role of observing responses in discrimination learning: Part II. Unpublished Ph.D. thesis, Univ. of Indiana, 1951.

(Received March 23, 1957)



## NEURAL MECHANISMS IN PERCEPTION<sup>1</sup>

JEROME S. BRUNER

*Harvard University*

I have been asked to offer a paper on perception, one that may be of some relevance to the manner in which the brain mediates perceptual processes. I admit to no special expertness in the neurophysiology of perception, although I have followed the growing literature with an increasing interest and a sense of excitement. There was a time, let me confess, when as a student of perception and thinking I operated on the working assumption that as far as I was concerned the nervous system was made of green cheese. This form of intellectual cheek, this know-nothing attitude, gives me little comfort nowadays. Yet many of my colleagues assure me that nothing we can find out about the nervous system can in any sense "explain" what we know about the subtle mixture of sensory, mnemonic, and inferential activity that we call perception, for, as they say, psychological theory does not reduce to neurophysiology any more than neurophysiology reduces to electromagnetic theory. I believe both of these assertions to be true. Yet, I have a lingering doubt. Indeed, I can even illustrate my doubt.

Some years ago I came to the conclusion, to simplify the matter grossly, that perceptual selectivity seemed to operate like a gating process that had the characteristic of raising the identification threshold for certain classes of stimuli and lowering them for others; and I suggested that rather than viewing such threshold differences as due to speed of

associational processes set up by different inputs, inputs might be conceived as being differentially gated out before ever reaching a stage of association. My argument was based on considerations of economy of functioning—that very likely perception had as one of its functions the task of screening out irrelevant stimuli and letting through relevant ones, an important consideration for living organisms who have a perilously narrow attention or immediate memory span. Implicitly, I was resisting a classical bit of neurophysiology as well: the doctrine about projection areas being linked up directly with association areas that enriched or "gave meaning" to sensory inputs. So I obviously was not completely independent of neurophysiological issues even at that time. During the past year, moreover, I have had occasion to read papers by Hernandez-Péon and associates (28) and by Galambos and associates (20)—both indicating that neurophysiologically, too, there seems to be evidence that screening or gating of inputs takes place as far peripherally in the auditory system as the cochlear nucleus. A cat, distracted by a mouse introduced into its visual field, ceased to show normal evoked potentials from this nucleus, far out though it was in the system. I was quite overjoyed—and in a manner not fit for somebody who claimed to believe that psychological theory must be self-sufficient—that it must not look for explanations in the black box of the nervous system.

Donald Hebb (24) provides me with an explanation of my unseemly joy. In a delightful paper in which he directs his shafts against the more unthinking type

<sup>1</sup> Presented at the 1957 meeting of the Association for Research in Nervous and Mental Disease. I am particularly indebted to Dr. Karl Pribram and Professor Walter Rosenblith for their helpful advice and criticism.

of antiphysiological bias in psychological theory he remarks, "The idea in rejecting physiology was to use only 'purely behavioral' conceptions, but some of these were actually of physiological origin and continue to exert a physiological influence on psychology" (24, p. 40). Or, as he says in a paper some years later, "... the psychologist who avoids physiological conceptions merely succeeds in avoiding modern ones, and is likely to have his thinking dominated by older ideas, vintage of 1890" (25, p. 404). And in fact, this has been the grave difficulty with physiological notions in psychology: their half life is too long, so long that their origins are forgotten and they become psychological gospel. My malaise with the associational theory of selective perception was a dissatisfaction with old neurophysiology masquerading as psychological theory. This is not to say that I wish to offer psychological theory to the contemporary neurophysiologists, but to find out, rather, what kinds of parallels and linkages might exist between what we believe to be the case in perception and what is gradually emerging concerning neurophysiological mechanisms that might mediate such processes. I might even hope that the way we as psychologists talk about perception might suggest some neurophysiological ideas to neurophysiologists!

### THREE FUNDAMENTAL PROBLEMS

There are three broad questions that one may ask about the perceptual process and these comprise, I think, most of the diverse activities that research workers concern themselves with in the study of perception; they are questions which, I think, lead most directly to neurophysiological speculation and inquiry.

The first of these is "How does perception represent the structure of the physical environment that constitutes stimulation?" Questions of this order

are familiar. How, to take an example from some work of Pfaffman (39), does salty taste get transmitted to the cortex in such a way that it can there be distinguished from quinine or acid? Or how do organisms learn to represent the probability structure of the environment, so that they may go swiftly and easily from a few cues like color, shape, and texture to the identification or inference of, say, an apple? There are far more complicated versions of this question, but these we can consider sometimes. Indeed, if there is anything to be said for the James-Lange theory of emotion, one must include here the question of how afferent impulses from the autonomic system are represented perceptually.

The second question is also simple—deceptively so, like the first one. "What accounts for the fact that perception may remain relatively constant while the physical stimuli operating at the receptor surface are varying drastically?" Here we are dealing with several seemingly different kinds of phenomena. The first is the phenomenon of constancy: the size, shape, color, and brightness constancies. In a field with distance cues available, an object moving away from the observer does not change in apparent size at anywhere near the rate required by the inverse square law that governs the size of the retinal image. Or, with field-illumination cues present, two surfaces with the same albedo appear of about the same brightness, even though one be in shadow and the other in diffused sunlight and their total luminous flux different by an order of magnitude. Moreover, for all save the very young child, objects maintain their identity though they change appearance—seen from the front or the rear, disappearing behind a screen and reappearing, going through various distortions. A range of highly variable figures are perceived and discriminated as triangles so



long as they possess certain criterial properties and regardless of other varying features, such as difference in size, texture, and so on. Indeed, the phenomena of linguistics are rife with examples of such constancy in the face of perceptual variability. Wide ranges of speech sounds are rendered constant, if they comprise the range included in a phoneme, and it is characteristic that variations within a phoneme are not noticed. Given the phonemic structure of Japanese, the naive Japanese listener does not notice the difference between "rice" and "lice," "lobster" and "robster"—the latter word having actually appeared in a Japanese-English magazine in place of the former.

The third question covers an equally wide range of problems. "What accounts for the fact that perception may vary while the stimulus input remains constant?" Here we deal with phenomena all the way from spontaneous alternation of such unstable three-dimensional figures as the Necker cube to reorganization of the perceptual field under conditions of change of set. One moment we do not understand a sentence spoken to us and the next moment we do, though it has been repeated in identical form.

#### PERCEPTUAL REPRESENTATION

When we say that perception represents the physical world of stimuli with which an organism has sensory commerce, we mean essentially that perception is varyingly predictive. That is to say, to take a banal example, if one stick appears longer than another when some distance removed from each other, they will likely show the property that one will overlap the other when placed one next the other, or that one will actually "measure" longer than the other when the technological aid of a foot rule is employed.

It was the objective of Weber (54),

then of Fechner (13), to state a general psychophysical law to describe the translation of stimulus magnitudes into subjective magnitudes. The most recent reformulation of this general law is Stevens's: "... there is a general psychophysical law relating subjective magnitude to stimulus magnitude, ... this law is simply that equal stimulus ratios produce equal subjective ratios" (46, p. 153). We shall return to this matter shortly.

Even when such simple prediction fails, the malprediction is more often than not systematic, and still contains the basis of a prediction. Add to the sticks the tails required to create the Müller-Lyer illusion and the simple foot-rule prediction breaks down; but the breakdown is systematic and, if the perceiver has the right transformation equation with which to correct his prediction, he will still come out all right. This is not to say that perceptual experience is a unique isomorph of the physical attributes of the stimulating environment, for patently it is not. There is error of a random type also operating. Partly the error can be thought of as a result of the imperfect resolving power of the sensory receptors and their related neural tracts and centers. Partly, the error can be attributed to the presence of various forms of centrifugally generated noise, or even, possibly, to the operation of centrifugal control mechanisms of the type discussed by Granit (22).

There is another form of representation which, on the face of it at least, seems of a different order. It has to do with the representation of *objects* in the environment, and such representation must obviously be learned. It consists, to use the phrase of the brilliant, and tragically dead, young Cambridge psychologist-physiologist, Kenneth Craik, in a process of model building, the construction of neural models



or "templates" to represent the redundant structures of the environment. Nowhere is the importance for perception of such representing models more evident than in the study of perception under tachistoscopic presentation. Given a brief flash of a complex object, say a picture of a human figure at 10 milliseconds in good illumination, a subject picks up a few defining cues, infers the rest, and with little difficulty and considerable confidence announces that he has seen a man, a generic man, to be sure, but a man nonetheless. If what has been presented in the tachistoscope reflects the high-probability redundancies of the environment, our subject will be right. If, however, we make up the stimulus picture with the head of a giraffe, the arms of a gibbon, the trunk of a bear, etc., then our subject will be, for practical purposes, the victim of his redundancy-matching models. We know from the brilliant observational studies of Piaget, particularly from his *Construction of Reality in the Child* (40), and from the somewhat ambiguous accounts of von Senden (44), the difficulties encountered in perceptual identification when there has been no intervening opportunity for an organism to construct such representing models for environmental objects.

Let me recount briefly some of the characteristics of perceptual representations as these have come to be known over the last half-century of relatively uninterrupted research on this topic.

1. *The categorial nature of representation.* What is most characteristic of perceptual representation is that it is categorial in nature, better described in the language of set theory and Boolean algebra than any other. An input is allocated to a class of objects and achieves its identity thereby. We see triangles, lines, apples, people. Ordering by subjective magnitude can be shown to be subservient to this first primitive,

rather conceptual placement. "Big" or "small" as a subjective judgment depends upon the summary midpoint of the class in which the object has been placed, and it is still something of a moot point how the nervous system estimates the first moment of magnitude or quality of the distribution of objects that are arrayed in a class. It is further the case that the process of achieving placement of an input is often an iterative estimation procedure.<sup>2</sup> You are trying to guess the square root of a number, say 4. Guess any number less than 4. Divide 4 by this number. If the quotient is bigger than the number you divided by, the guess is too small, and vice versa. So now take the mean of the guess and the quotient. Use this as your next guess. In a few turns, you will be very close indeed to the square root. So it seems to go, in a general way, with categorization of inputs. There is a first guess—determined by the expectancy of the organism. It is tried out and found to "fit" only approximately. The result of the attempted placement is then somehow averaged with the original guess, and on it goes. This is a gross oversimplification, but it will serve us for now. What it fails to describe is the frequent discontinuity of the process of placing, but we will come to that when we speak of the variability of perception in the face of a constant stimulus.

The iterative or, more likely, Markovian nature of successive approximation in perceptual identification also makes more understandable both the systematic nature of errors in identifica-

<sup>2</sup> I am grateful to Professor R. C. Oldfield of Oxford for bringing this possibility to my attention. The analogy and disanalogy between iterative procedure in mathematical estimation and in cognitive operations is developed in his "Some Principles of Convergence in Thinking," presented to a Symposium on Thinking held at St. John's College, Cambridge, during the summer of 1955.

tion observed in tachistoscopic perception and the abrupt manner in which a long series of increasing errors sometimes gets corrected. If, for example, a subject is presented successively at rapid exposure with a stimulus array containing an incongruous element—say, a picture of a discus thrower, wound for the throw but holding a bass viol in his free hand—there will be a series of iteration-like attempts to deal with the incongruity soon after the perceiver is able to report that he perceives an athlete in action. It is a “shadow across the front of the athlete” or “a track official kneeling” or a “sprinter going past on the track,” and these hypotheses become elaborated on successive exposures: inhomogeneities are seen as the track official’s hat or as a sweat suit on the passing sprinter. Finally, the elaboration reaches a point where its lack of fit become obvious and there is a notable disorganization of perception, until a new series of approximations begins (cf. Bruner and Postman, 7).

Perceptual identification or placement, then, can be likened to the determination of the fit between a model and some sample that is being matched to it. We know precious little about the nature of the models that are constructed neurally to represent the attributive and probabilistic structure of the individual’s environment, although various models have been proposed, depending in the main upon the development of reverberatory circuits and synaptic boutons, as in Hebb’s conception of cell assemblies and phase sequences (23), or upon reduplication patterns propagated across the entire cortex, as in Lashley’s ingenious conception of memory traces, concerning the failure of whose location he writes so provocatively (32). But it is not about these structures that I wish to speculate, for there has been enough speculation on this subject in ratio to the meager findings available. Rather,

it is with the process of matching or placing of inputs that I wish to concern myself briefly.

We know that rhythmic activity from the surface of the brain is related to attention, and that alpha rhythm is disrupted and replaced by more aperiodic rhythms when attention is directed to a stimulus. Let me put this finding together with two others. The first relates to the rhythm recorded from the olfactory bulb of the rabbit by Adrian (2), a rhythm of high frequency of 80–100 per second that continues so long as the animal is breathing regularly. If an odorous substance be introduced or the animal begins to sniff, “the rhythm disappears or gives place to the slower rhythm which occurs at each aspiration and is due to the olfactory discharge. Within a short time, 20 seconds or less, the rapid rhythm returns although the olfactory stimulation continues . . . but a change in the stimulus may suppress it again.” Adrian then goes on to say,

The organization of the olfactory pathways may be quite different from those for other forms of sensation, but it is natural to ask whether the rhythms found in the cerebral cortex may not operate in the same way, whether their function or part of it is not to block the passage of further afferent discharges when the information has served its purpose. The reticular formation might well be the decisive factor in the direction of attention, i.e., in the suppression of the rhythm in one region, for the signal arriving by the direct route to the cortex reinforced by one from the [reticular] formation might disrupt the rhythm and gain a clear path, whereas a signal not so reinforced would be unable to break the barrier (2, p. 243).

It should be noted immediately, however, that too much arousal reaction may have the effect of interfering with the signal arriving at the cortex. Psychophysical experiments show, on the whole, that too alert an observer pays a price in a raising of sensory thresholds, and that a certain amount of relaxation in the observer yields the lowest thresholds.



Yet, this special point aside, Adrian's observation fits well with what we know of the operation of attention.

A recent paper by Sharpless and Jasper (45) provides parallel and more elaborated evidence on the role of arousal reactions—evidence on the habituation of such reactions. These writers point out that it has been known for twenty years that with direct recording from the cortex of unanesthetized animals one may observe habituation to such ordinarily attention-compelling stimuli as sharp whistles and claps. The authors set themselves the task of analyzing the locus and the specificity of such habituation effects, recording from the cortex and from subcortical areas in cats with chronically implanted electrodes. Stimuli were repeatedly presented to normal sleeping cats whose electrocorticograms showed a clear sleep pattern. With habituation to a given stimulus, the duration of activation declined and its latency increased strikingly, sometimes showing a 20-second period between stimulation and activation onset. Eventually, complete habituation would occur. These patterns and the habituation effects were recorded from the cortex as well as from the mesial thalamus, the posterior hypothalamus, and the midbrain reticular formation. With respect to the specificity of habituation, a critical point in our discussion,

If either the intensity or duration of a stimulus to which the animal had become completely habituated was increased, it became capable of arousing the animal. More interesting, however, was the specificity of the habituation process to the quality of the repeated stimulus. Thus, if the animal had become completely habituated to a repeated tone, it could still be aroused by a light touch, a change in illumination, or by a tone of a different frequency (45, p. 662).

While the authors report some generalization of habituation to tones differing somewhat in frequency, they remark that

it is variable. They note also that "habituation was occasionally specific not only to the pitch of a sound, but also to the arrangement of tonal elements within the sound—its pattern." If a cat had been completely habituated to a tone falling in pitch from 5,000 to 200 c.p.s. in 4 sec., presentation of the stimulus pattern in reverse—rising from 200 to 5,000 in the same time—would produce activation. Destruction of the auditory cortex did not alter the pattern of habituation observed in intact animals nor its specificity—save for habituation to pattern, which was not found in such operated cats. Subcortical lesions, however, seriously disrupted specificity of habituation to tonal pitch, notably subcortical lesions that prevented direct passage of signals to the cortex but left intact the collateral pathways to the reticular system through which auditory signals could get to the hemispheres.

In sum, then, the evidence of Sharpless and Jasper seems to indicate that the arousal pattern that may subserve the selectivity of attention is capable of considerable selectivity of habituation to environmental stimulation. If habituation has this degree of specificity, it seems not unlikely that the obverse process, selective attention built up through training, might operate by a parallel subcortical mechanism—selectively reinforcing not only "non-habitual" stimuli but also stimuli to which, through training, special significance had become attached. Such reinforcement may be essential for the stimulus input to reach a level of activation where a "fit" can be made between an input and a "model."

The test of this is in principle not difficult to make although it is probably very difficult technically—at least for the while. If it can be shown that there is an onset in arousal impulses from the reticular system that occurs while an observer is trying to "figure out" the nature of an input, and a dimi-



nution following the identification of the stimulating object, then it would seem more plausible still that "effort after meaning," to use Bartlett's phrase, is dependent for its fruition upon reinforcing stimulation from the reticular system—either to clear the track for it, or to make possible the process of matching.

2. *The summary nature of perception.* The act of perceiving almost always is accompanied by a loss of information. That is to say, if we think of the perceptual process as an input-output system, we find that no matter how we measure the nature of the input—either by physical measurement or by a consensus of observers—the limited attention span of the human organism (and presumably lower organisms as well) imposes loss in the transmission of information, even assuming that internally generated noise could be eliminated. This seems to me to be intuitively obvious, though it ends one in much complication trying to measure the nature or extent of the loss. The only known way of combatting loss in transmission is by a process of what Miller (37) has recently described as "recoding of input." If it is the case, and the evidence speaks rather overwhelmingly for it, that attention or immediate memory span is for about seven units plus or minus two, the recoding process consists of regrouping the input such that each of the seven places, so to speak, is filled with more valuable coin. A good example is the phenomenon of segmentation in language perception—perception either of speech or of written language. Presented with an array of letters, we do not try to take in the individual letters, but to recognize words. While we have a span of from eight to ten random letters, the span for letters when grouped into words is much greater. So too with the speech flow and its segmentation into words on the basis of phonemes. But no recoding is so perfect that it

makes it possible to scan a great slice of the universe and recode it without information loss, for most of the forms of recoding we employ fall considerably short of such comprehensiveness, so that wherever we turn there is information loss.

Now may I quote a debate reproduced in the volume *Brain Mechanisms and Consciousness*, the symposium on the reticular system held in 1953, a debate between Bremer and Fessard. Bremer is criticizing the idea that integration of perception occurs at a locus; the particular locus he is decrying is the reticular formation. "I would certainly not have chosen the reticular formation for that location on account of the fact that information is lost there by reason of neuron convergence and occlusion. Conscious integration excludes dilution and loss of information" (5, p. 245). Fessard defends his conception of locus, and particularly the locus of the reticular formation, in part in these words, with which I thoroughly agree without taking sides on either the notion of a specific locus or the role of the reticular system:

Dr. Bremer has judiciously pointed to its main difficulty, that is, to the loss of information that would result from the simultaneous arrival of impulses from different origins upon the same neurone, as we now know that this happens in the reticular formation. I admit that this difficulty is not easily disposed of, but I do not agree with Dr. Bremer when he says that integration cannot go with loss of information: psychologists would certainly support the idea that integration is only obtained at the cost of big losses of detail, as when a perceptive Gestalt is apprehended as a whole, its components being totally ignored (14, p. 248).

I should like to propose, as an hypothesis, that both Bremer and Fessard are correct—the former in denying locus to integration, the latter in assuming that loss through convergence and occlusion is a feature of afferent transmission, just as loss of information is a fea-

ture of perceiving generally. Two lines of argument and evidence seem to recommend themselves here. The first is Lashley's argument about summation in the afferent system; the second is the point of view put forward by Rosenblith (42) in discussing the bioelectric fate of a click from the round window to the auditory receiving area.

Lashley, as you know, finds the summative characteristic of neural activity a necessary parallel to the nature of conscious experience itself.

A point of fundamental importance for a theory of the neural basis of perception is that there is never awareness of the integrative activity of the brain while it is in progress. The perceived items are always the product of preceding and complex integrative processes. Visual distance is a good illustration of this. Things are seen as near or far, yet this distance is actually determined by a number of variables, binocular parallax, estimates of relative size, texture, etc. which are not separately perceived but are only revealed by experimental isolation. . . . So in every case, that of which we are aware is an organized structure; the organizing is never experienced (33, p. 424).

Lashley then goes on to consider neural summation.

Consideration of the role of summation in neural activity reveals a similar characteristic. Even at spinal levels summation of excitations is essential for reaction. Summation combines a multiplicity of excitations into a single effective unit which the resultant reaction cannot analyze into separate elements. So also we must conceive of higher level integration as a sequence of excitations, individually incapable of exciting further activity, but continually summing to arouse the succeeding neural patterns (33, p. 424).

It is perfectly apparent that given such a model of summation—and I shall assume that it inevitably involves stages of convergence and occlusion—it is not necessary to suppose that integration has a particular locus. It is the entire range of summative activity, from the brain stem and lower on up to the cortex, that enters into integrative action. I will

make bold to propose that an understanding of the *nature* of this summative activity is what will provide the key to an adequate understanding of how it is that perception loses information both systematically and randomly, and thereby manages to achieve integration and summary representation.

Let me turn briefly to one proposal concerning the nature of summative activity, to a model inherent in Rosenblith's conception of click transmission in the auditory system.

We are inclined to interpret the electrical responses that our wire electrodes record as summated activity from a sample of a population of neural elements at a given station of the auditory nervous system. The number of neural elements that contribute to a response depends upon the intensity of the stimulus that is used. As has been shown [by McGill and Rosenblith (36)], the amplitude of an evoked response can—within certain limits—be interpreted as an estimate of the probability of response of the neural population under investigation (42, p. 5).

Rosenblith is concerned in this passage principally with the problem of amplitude of response, and it may well be that the statistical operation of the auditory and other systems works by a simple system of unbiased sampling-summation where amplitude is concerned. But there is also good evidence that higher-level units respond to biases in the sample of lower-level units that summate to activate them. The best evidence of this is provided by close analysis of the systematic patterns of discharge that occur at lower levels in response to known physical properties of stimulation. Again audition serves us well. Consider the second-order neurones in this system, as studied with microelectrodes by Galambos and Davis (18, 19). As Davis summarizes it:

Each second-order neuron is "tuned" to a particular frequency. Less acoustic energy is required at this frequency than at any other to set up impulses. Near threshold, the tun-



ing is very sharp, but the neuron responds to a broader and broader band of frequencies as the intensity is increased. . . . Even for fairly strong tones the tuning of each neuron to a particular frequency is still apparent (9, pp. 1128-1129).

This is the situation at the second-order neurones, after transmission across a single synapse. Unfortunately, the data available on the pattern of firing in the spiraling single neurones of the short eighth nerve (47) do not provide a simple picture of how firing patterns in the first neurone signal across the synapse to the second neurones of the auditory system, nor do they indicate whether there is a summation pattern operative there. When we tune the organism to listen to pitch or to loudness or to density of a tone, or to some particular pitch or loudness or density, it may well be that the nature of the summation required for firing a higher-order unit changes as far down the pathway as the first synapse, and that the resultant final message to the auditory receiving area reflects these changes in nature of summation. We must beware of being too impressed by the elegant and relatively invariant cortical mapping of auditory frequencies found for the anesthetized dog's brain by Tunturi (50, 51). Anesthesia plays funny tricks and, as Békésy (4) has suggested, "for every level of anesthesia we will probably have a hearing theory."

What I would like to suggest—and I have no neurophysiological warrant for doing so but only a hunch based on psychological observation—is that the form of biased sampling of lower-order units that determines the summated activation of higher-order units is very likely controlled by centrifugal control mechanisms that serve to program lower-order stations in the afferent system. These may be thought of as orders for types of summation that can effectively trip off higher order neurones on which lower

order neurones converge, and it may well be that such changes in nature of summation account for the way in which the summary nature of perceiving changes as a function of change of set in the organism. Adrian puts the matter metaphorically:

. . . in normal waking life the picture that is presented to consciousness has many details left out. First one element of the pattern claims our attention, then another. At some stage the complete report from the sense organs must be subjected to an editing which emphasizes the important items and sets the unimportant aside. There is clearly some such editing at work in most of the sensory pathways. . . . It might be expected wherever the signals pass through a sheet of closely connected neurones (2, p. 239).

He then goes on to suggest two possibilities as to where this editing might occur:

One is that the afferent signals from, e.g., the ear may be summarized on their way up the auditory pathway, but that the content of the message is allowed to reach the cortex whatever may be going on in the other sensory pathways. This means that at any moment the cortex would have at its disposal the complete pattern of the environment provided by the sense organs, but that there is some controlling mechanism (of which the reticular formation may be a part) to decide that the auditory elements of the pattern should have priority and to make the auditory region of the cortex specially receptive and the other areas less receptive. The other possibility is that the controlling mechanism operates at an earlier stage and that we attend to the sound and are unaware of contacts and pressures because the auditory signals are allowed to reach the cortex and the tactile signals are not (2, pp. 239-240).

Is it not conceivable that successive "synaptic editing" could occur, controlled by centrally induced programming of summation characteristics, and that there could also be the kind of "boosting" operation discussed earlier wherein a system like the reticular formation reinforces the edited pattern of stimulation that reaches the brain?

To sum up, then, we have proposed



that information loss in perception can be viewed as the resultant of successive summation at stations along the afferent pathways, that this summation is likely statistical in its characteristics, that the sampling of lower-order units that summate to fire higher order units is biased by a gating or programming of way stations along the pathway controlled by more central mechanisms, and that the resultant in perception is a biased summary of the external stimulating environment. The biasing is reflected in the systematic errors of perception and in the forms of selectivity that make it possible for perception to vary in the face of a constant pattern of external stimulation.

3. *The autogenic properties of representation.* If it can be said that the representation of the external stimulus environment is categorial and summary in nature, we must add to these properties that representations, once accomplished, appear to generate processes of their own. Here we must consider briefly some of the properties of the models that organisms construct to represent their environment and its objects. It was suggested some years ago by Craik (8) that neural models, once established—whether in the form of traces, schemata, cell assemblies, or whatever short-hand terms one uses—have the property of making imagery possible in the absence of stimulation. A recent work by Mandler (35) on the “learning of imagery” suggests that some such function is served by learned structures. His experimental procedure is very simple. The subject has the task of learning the correct pattern amongst a set of spatially arrayed levers on a board. The order of correct levers is determined at random. In a given number of trials, the subject masters the correct order in which switches are to be thrown; he is able to perform at a criterion of 100% correct responding. The trials are now continued and

overlearning carried many trials beyond the number needed for mastery of the task. After a certain amount of overlearning, subjects begin to report not only that they can choose the right path but that they have a spatial image of the order, a spatial image that they can keep in mind even when the lever board is not before them. It is tempting to speculate what this means. When neural integration has progressed to a certain level, the organization achieved appears to be able to generate mental content of its own. I do not pretend to understand what kinds of neural mechanisms are at work here, but I commend the problem to you as perhaps worth pursuit by the use of electroencephalographic methods. When learning has progressed to the stage of overlearning and imagery production, are there detectable changes in the pattern of EEG to be recorded? Methods for integrating and analyzing complex wave forms now widely in use in work on EEG might cast some first light on this problem that would permit more enlightened next steps.

There are other autogenic properties of what I prefer to call operative cognitive structures that are worth a few words of description on the chance that they might make contact with work now going on in neurophysiology. The first of these is a *normalizing tendency* exerted by cognitive models on the percepts they generate when stimulus input occurs. Let me first give an example and then explain what I mean by this process, usually called assimilation. For each category or model constructed to represent objects or events in the environment, there appears to be a set of average property values characterizing the typical instance of the class. It may be likened to the adaptation level in psychophysical judgment about which Helson (26) and others have been concerned. It is a reference point in the class which, in judgments of brightness

or weight, is best estimated by computing a weighted geometric mean of instances previously encountered. Subjective magnitudes greater than this midpoint are described as "heavy" or "bright," smaller as "light" or "dark." With respect to more complex objects, say apples, there are a set of such specifications of a typical apple—a modal color, shape, size, texture, etc. In the event of masking noise or of ambiguity in the nature of a stimulus input, one finds normalization of the percept taking place by assimilation of aberrant or ambiguous attributes to the modal values of the category where the input has been coded. Or, in the case of inattention to particular attributes, modal values of the category will be substituted for what has not been noticed. Such normalization or completion is one of the principal sources of error in perception—error in the sense of failure of precise representation. An average is substituted for a specific value. The normalizing process continues, moreover, in the process of forgetting—forgetting being more often describable as regression away from specific characteristics of stimulus event in the direction of the modal characteristics of the category into which the stimulus event has been placed.

I readily recognize that neurophysiological research has not proceeded far enough to have many sound conjectures about the kinds of structures that can produce the types of autogenic activities just described. Yet it is well to bear such characteristics in mind in the process of setting forth hypotheses about the structures. Looking at the field from the outside, I have the impression that the neural mechanisms proposed, even in a half-playful spirit, tend to be built for the purpose of handling a single psychological phenomenon. There are too many plausible neural models of the one-phenomenon type. It is only when one

requires of them that they accomplish several forms of psychological resultant that the range of possibilities narrows realistically.

#### PERCEPTUAL INVARIANCE AND VARIABLE STIMULATION

Let me begin with Fletcher's well-known study of speech perception (16), a striking example of what Lashley has for a long time called "the equivalence problem." A sample of speech is presented, treated in such a way that all frequencies *below* a certain cutoff are filtered out. Now the same speech sample is presented again, this time all frequencies *above* the cutoff being filtered out. In both cases, the speech pattern is 100% intelligible to listeners. Frequency, we then say, is not a criterial attribute of speech; one band of frequencies will do as well as another. And so it will go with most forms of perceptual identification, given, as we have noted, its categorial nature. The task in any given study of perceptual equivalence is to determine by reduction methods what particular stimulus attributes are criterial for identification of an object as belonging to a certain class or, more simply, as being a certain thing. Formally speaking, the determination of the criterial attributes of an equivalence class is the same whether one is at the perceptual level or at the level of providing the defining properties of such an abstract concept as force or momentum.<sup>3</sup> My concern here is to consider what is involved in the more perceptual forms of such activity.

Virtually every writer who has tried his hand at a neural theory of perception has been mindful of the problem; it is not overlooked. Adrian perhaps puts the matter most succinctly:

<sup>3</sup> For the reader interested in the continuities in all such categorizing I would suggest the work of Bruner, Goodnow, and Austin on categorizing and thinking (6).



We can recognize a tune whether the pitch is high or low, and we can recognize a triangle or a letter, whatever its size or whatever the position of the image on the retina. Now this must mean that there is no need for particular sensory endings or particular pathways to the cortex to be excited; for the recognition of the pattern all that seems to matter is that the excited regions should have some general relations to one another. There must be something corresponding to the three angles of the triangle, but there is no need for each of them to be represented at a particular point in the receiving area of the cortex (1, pp. 81-82).

There has been much bitter ink spilled on the question of equivalence, notably amongst psychologists who tend to take one of two basic positions. One of them, stemming I think from neurological theories of a half-century ago, is that equivalence occurs by the formation of a generalization gradient, a kind of irradiation from a particular stimulus value. The model is familiar from Pavlov's work on generalization, and it is this model that has dominated much thinking in American psychology. The organism learns to respond to a given value of a stimulus or to a given object plus a range around it. Lashley (33) has been particularly forthright in attacking this view, noting that the model probably holds only for instances of pseudogeneralization where there is failure to discriminate the original stimulus from one presented later. There is much that is logically awkward about the simple generalization notion that various writers have tried to patch up in one way or another—some quite plausibly. It does not explain how certain properties come to be the defining properties of a class and others irrelevant. To handle this, a reinforcement theory is invoked, and the defining properties are said to be determined by the composite attributes that are always present when reinforcement occurs, the remainder averaging out in strength because of their random relation to reinforcement. But, as Woodworth (55) has wisely noted, what can we take as

reinforcement in perception? In most theories it is taken to be something that reduces the need of an organism, related to attaining a goal. To extend this to the question of learning the properties of the class "triangle" is somewhat difficult. In any case, since the controversy does not concern us directly, we can leave the matter with the statement that the neural model of a class assumed to be constructed in this way is a bit like a composite photograph in which the tried and true elements remain, and the others are blurred out. Such a model of integration, suitable though it may be for describing how it is that the three angles emerge as the diacritica of triangles while size and texture and relative length of sides are blurred out, fails to deal with the essentially transformable properties of groupings once they have been formed; nor does it show how it is that in one moment of insight one comes to group all plane figures as resolvable into combinations of triangular forms and treats all plane figures as members of a new class of simple and composite triangles.

It seems to me that the problem of equivalence, leaving aside now how the primitive equivalence category is formed, is essentially the problem of how an organism processes a stimulus ensemble, what features of input it is responding to. It is only with respect to this program for processing environmental stimulation that it is possible to talk about an adequate stimulus or a just noticeable difference in stimulation. Almost invariably, as both Adrian (1) and Lashley (33) have suggested, stimulus information is processed in terms of *relationships* either in a spatial array or in a temporal array, and it seems highly likely that spatial relations are translated into temporal ones in the nervous system—given the requirement of time for scanning a spatially extended field. What this suggests, then, is that the



equivalence of stimulus events is a function of certain invariances in relationship in a temporally extended neural excitation; certain patterns of change in excitation define equivalence.

There is very suggestive evidence concerning the importance of temporal change as providing the pattern on which equivalence is based. A famous passage of Exner's (12), where he compares the scanning movements of the copepod *Copilia* with the scanning movements of the eye, comes to mind, "Wir tasten mit unserem Blick des Seefeld ab," and indeed it is this "fingering over" the visual field that suggests the importance of temporal integration. Thorpe (48), in his recent and most excellent book on animal behavior, gives various examples of the importance of temporally extended scanning of a stimulus array in animal discrimination. Dethier (10), for example, has shown that the caterpillars of *Lepidoptera* can respond to form and form equivalence in spite of the optically inadequate stemmata placed on either side of the head, optically insufficient in resolving power for representation of form. Yet form can be perceived by these organisms provided it is possible for them to move their heads around in relation to the object to be discriminated, by the familiar waving or "inching" movements of the anterior part of the body of these caterpillars. Some sort of scanning process, extended over time, results in a temporal summation and/or integration that makes form discrimination possible. So, too, in the work of Ulflyott (52) on klinokinesis and klinotaxis in the flatworm *Dendrocoelum*. Again, a primitive optic apparatus is involved, supported either by the animal's locomoting by a convoluted path with a relatively constant rate of change in direction (klinokinesis), or by proceeding in a relatively straight line with rather constant waving of the an-

terior section of the body where the sense receptor is found (klinotaxis).

Given the nature of adaptation of sense organs, it is not unlikely that equivalence coding almost requires the use of patterns of change in excitation over time. Not only do we have the facts of adaptation with respect to intensity of stimulation in vision, somesthesia, etc., but there is now accumulating evidence from the work of Riggs (41) and of Ditchburn (11) that if change of pattern on a sensory surface does not occur at a quite rapid rate, perception of form and detail adapts out. Ditchburn, using an ingenious mirror attachment on the cornea, casts a simple form image on to it, collects the image, and directs the beam back into the eye. The arrangement keeps the cast image at a constant point on the retina, resting on this point in spite of eye tremor. Within six seconds, the simple image, a fine-lined reticule, disappears. The moment one jars the apparatus, even slightly, the image changes position however slightly on the retina and reappears. These results leave little question that change over time is needed even for the maintenance of a simple image.

I should like to suggest that equivalence of an array of stimuli is essentially a centrally determined program concerning the acceptance limits of temporally extended change patterns. Where and how these patterns are acted upon and coded for equivalence, I have no clue. Adrian and Moruzzi (3) show suggestive evidence that the electrical pattern in the somatic cortex is much alike when the same place on the foot is touched, so that it is possible that the cortex has enough information to do all the job of coding for equivalence in terms of some program of temporal integration.

Let me say, finally, in support of this general view, that most of the findings in the constancies seem to support the

view that apparent brightness, color, speed, and form—all of them exhibiting notable tendencies toward high perceptual invariance in the presence of marked change in the physical stimulus—turn out to be explicable in terms of constant relationships existing in the stimulus ensemble. Wallach (53) has shown that, within wide limits, apparent brightness is a function of ratios of stimulation in the visual field. A central field surrounded by an annulus with a ratio of 1:10 in luminous flux appears to be the same brightness as a central field that bears a relation of 10:100 to its annulus. Two objects are seen as traveling at the same speed if their angular displacement bears a similar relationship to their apparent distance from the eye—a more complicated form of ratio. And as Gibson (21) has shown, apparent distance is a function of the change in texture of the field intervening between the object and the eye. He distinguishes four temporal or spatial orders: Let *l* and *d* indicate light and dark. Then the order *llll* or *dddd* will be seen as a real or film color; *llddd* or *dddll* as a line or the edge of a figure; *ldldld* as a surface; and *lllddddlllddddldld* as a gradient of texture indicating distance, equal distances presumably being based on like gradients of texture, all other things being equal.

Undoubtedly, some of the equivalence coding in the perceptual constancies is innate, some learned. When we come to identity constancies, treating as equivalent objects that have been altered drastically in all respects save their defining attributes, we are very likely dealing with learned behavior exclusively. And here the "rules of equivalence" that are learned doubtless reflect the adjustment requirements of the organisms. We know from the work of Thouless (49), for example, that mechanics must and do learn to *depend* on size-constancy effects in judging the sizes of nuts and

bolts and the like, whether the objects are held in hand or are seen on the bench at some remove; and that painters learn to *overcome* brightness and color constancy effects to get on with their jobs. Klüver's brilliant monograph on equivalence of stimuli (30) indicates the degree to which equivalence grouping is generally subject to learning. But perhaps the matter will be better elucidated if we turn now to the problem of perceptual variation in the face of constant stimulation.

#### PERCEPTUAL VARIABILITY AND STIMULUS CONSTANCY

Let me begin by telling you of a game my colleague George Miller and I played at lunch one day. The question was whether one could construct a sentence pun rather than a word pun. Our two successes serve to underline the nature of selective organization of a flow of information. The two sentence puns are, first,

The good candy came any way.  
The good can decay many ways.

and

One's own leeway is to earn ever.  
One's only way is to err never.

Spoken rapidly, each speech flow in a pair is virtually identical from a phonemic point of view. In the first pair, a final *s* provides the cue for retroactive segmentation of the flow into lexemes or words. In the second pair, there is no such cue provided. The perception of speech is full of such examples of differential ways of organizing a temporal flow of stimulation. Since the pioneering work of de Saussure (43) on the isolation of the phoneme, and with the development of Jakobson and Halle's method (29) of decomposing the phoneme into a set of distinctive features, we know that the process of understanding speech involves a highly selective



method of isolating certain ranges of speech sounds, treating these quite arbitrarily as equivalent, and then using these as the diacritica by which words may be distinguished. There is often ambiguity in the process of segmentation, but this is rarely serious, for context almost always settles the issue—although the low-comedy deaf clown of classic vaudeville testifies to how close a thing such settlement can be.

It is a commonplace of psychological research that the organization of a complex perception can be varied by varying the set of the person—by varying the thing for which he is looking. Or, to use another common form of discourse, what we take in depends upon how our attention is directed.

What are the mechanisms that determine whether or not a given stimulus is perceived? What determines its dominance over other stimuli? Very intense stimuli generally dominate, and certain sensory modalities seem generally to be prepotent. This is particularly true of pain, perhaps of movement in the peripheral retina, of high pitched sounds, of some odors. In birds and higher primates vision is prepotent but probably not in other mammalian orders, though the evidence is uncertain. However, for man, and probably for most animals, the characteristics of the stimulus are generally of secondary importance in determining what is perceived. Internal factors, characterized as voluntary attention, play the major role in selecting stimuli for perception. Sometimes such attention involves chiefly giving prepotence to a sensory modality, as when one looks at a picture or listens to music. Even in such cases, however, there is a further selection of the items perceived, in terms of organization that is somehow inherent in the neural processes (33, pp. 426-427).

I think Lashley's summary of the matter admirable. What manner of neural processes seem to be involved?

Adrian proposes a first step toward an answer.

The operations of the brain seem to be related to particular fields of sensory information which vary from moment to moment with the shifts of our attention. The signals from

the sense organs must be treated differently when we attend to them and when we do not, and if we could decide how and where the divergence arises, we should be nearer to understanding how the level of consciousness is reached. The question [is] whether the afferent messages that evoke sensations are allowed at all times to reach the cerebral cortex or are sometimes blocked at a lower level. Clearly we can reduce the inflow from the sense organs as we do by closing the eyes and relaxing the muscles when we wish to sleep and it is quite probable that the sensitivity of some of the sense organs can be directly influenced by the central nervous system. But even in deep sleep or coma there is no reason to believe that sensory messages no longer reach the central nervous system. At some stage therefore on their passage to consciousness the messages meet with barriers that are sometimes open and sometimes closed. Where are these barriers, in the cortex, the brainstem, or elsewhere? (2, pp. 238-239).

I would like to suggest as a first hypothesis that the barriers are everywhere. The recent experiments of Galambos *et al.* (20) and Hernandez-Péon *et al.* (28) suggest that in the auditory system the barriers may exist as far peripherally as the second synapse at the cochlear nucleus. In the first of these studies it was shown that evoked potentials in the cochlear nucleus produced by click stimulation adapted out with repetition. The click was then paired with an electric shock delivered across the chests of the experimental cats, and conditioning established; following this, the conditioned reaction was extinguished and then re-established, etc. While the click was a signal for shock, there was greater electrical discharge from the cochlear nucleus as well as from higher centers—the caudate nucleus and the auditory cortex. When extinction occurred, when the animals did not show the crouching-snarling response to the click, there was marked reduction of activity in all three loci. To check whether the motor activity associated with the conditioned response was responsible for these changes in evoked potentials, Ga-



lambos and his colleagues checked their findings on cats in whom a complete muscular paralysis had been induced. No difference in the general findings was produced by this precautionary control.

The experiments of Hernandez-Péon *et al.* were of the same order, save that the attentional factor was varied in a somewhat different way. Again, activity was recorded at the cochlear nucleus, involving either second- or third-order neurones, and cats were exposed to pure tones at intensities comfortable to the human ear. Electrical activity was studied with three forms of distraction outside the auditory sphere: while exposed to the tone, cats were shown two mice in a bell jar, were given fish odors to smell, or were shocked on the forepaw; i.e., visual, auditory, and somatic distractors were employed. Electrical activity at the cochlear nucleus was markedly reduced under all three conditions in comparison with control responses. The authors propose that the suppression of the response in the cochlear nucleus is probably effected by inhibitory impulses from the midbrain reticular system.

The question arises whether the inhibition and facilitation effects found in these two studies are selective within a modality, whether there is selectivity for particular tones or other stimuli. The evidence here is still sketchy. Hernandez-Péon and Scherrer (27) have reported, however, that selective adaptation or habituation effects, known to occur in the midbrain reticular formation, can also be shown to operate at the cochlear nucleus. Prolonged repetition of a tone of a given frequency produces habituation in this nucleus, but if the frequency of the stimulating tone is changed, the electrical activity in the nucleus increases to normal levels. What remains to be done is to repeat the Galambos experiments with changes in the characteristic of the stimulating tone.

If the conditioned stimulus is a tone of 3000 c.p.s., will a tone of 5000 c.p.s. produce less electrical discharge in the cochlear nucleus, the caudate, and the auditory cortex?

Space does not permit a proper appraisal of the evidence for the assertion that selective barriers may operate in altering perception at virtually any level of the nervous system. Granit (22) presents evidence that suggests such a widespread locus of gating, inhibition, and facilitation. The work of Leksell (34) and of Kuffler and Hunt (31) on the control of afferent impulses from muscle spindles by the small diameter fibers of the ventral root of the spinal cord suggests that gating and control of impulses can occur right at the sense receptor. I do not know whether those of you who are expert in such matters would urge that, if there is such widespread monitoring of incoming impulses, it can all be traced back to the action of structures in what Penfield (38) calls the "centrencephalic system." I know that Professor Lashley's is a dissenting voice (33) to what appears to be a growing chorus of assent. Pribram<sup>4</sup> is also among those who would look for a broader locus, including a considerable role for what he speaks of as the posterior intrinsic sector of the dorsolateral parietal convexity of the cortex. He distinguishes in perception between "detection" and "identification," the former being a judgment of presence-absence of a stimulus and the latter being its placement or recognition. Destruction of the posterior intrinsic sector in monkeys appears to destroy their capacity for choice discrimination problems while leaving seemingly intact their capacity to respond to presence and absence of a stimulus. But it is interesting to note that Pribram describes the posterior intrinsic sector as receiving its extratel-

<sup>4</sup>Pribram K. The neurology of thinking. (In preparation.)

encephalic input exclusively from the thalamus, and it may well be that here again we are dealing with a central alerting or reinforcing station in the manner of Penfield's "centrencephalic system."<sup>5</sup>

But these are surely not questions of detail to be settled by a psychologist not properly cognizant of the anatomical and electrophysiological subtleties involved in locating where gating mechanisms are operative. My only reaction is that what is psychologically characteristic of selectivity on the afferent side is that one finds it at all levels of functioning—in the reception of information, in the selective losses of memory, and in the transformations of information involved in the thought processes. It seems highly unlikely, given this widespread range of selective loss and transformation, that there is one "seat" of this function in the nervous system, although it does not rule out the possibility of a centralized coordinating sector. Indeed, Galambos (17) has found not only that gating extends down to the first neuron of the auditory nerve, but that one can create gating at this level by artificial stimulation of the olivo-cochlear bundle, and he also shows that section of this bundle prevents such centrifugal gating. Since this has been demonstrated, one wonders

<sup>5</sup> Since the above discussion was written, the writer has been informed by F. Bremer of Brussels of an experiment carried out in his laboratory by M. Jouvét and J. E. Desmedt in which the electrical activity of the cochlear nucleus in response to a click is inhibited by concurrent electrical stimulation of the lateral part of the anterior mesencephalic tegmentum. The inhibitory control exercised by this mesencephalic locus does not appear to be mediated by the tract joining the superior olive and the cochlear nucleus. In any case it is becoming increasingly apparent that there are likely to be many centrifugal routes for the control, monitoring, and regulation of afferent input. A brief report of the work of Jouvét and Desmedt is to be found in *Comptes rendus des seances de l'Academie des Sciences (Belgian)*, 243, 1916-1917, seance de 5 decembre 1956.

whether we are not on the verge of a new era in neurophysiological research where the old model of the reflex arc is once and for all going to be dethroned. The model that is emerging is certainly one that accords more with the known variability of perception under conditions of constant stimulation.

Finally, let me say a few words about the nature of variations in perceptual organization as these occur in the presence of a constant or steady stimulus. In the preceding section, mention was made of a set of rules governing the manner in which the relations within a stimulus ensemble might be used for coding the input, and that so long as inputs showed the prescribed relationships in the patterns of excitation they elicited, they would be coded equivalently. The complementary point can be made about variability in perception with constant input. Given a change in the rules for the utilization of stimulus relationships, the resulting percept can change even though the stimulus remains the same. Pribram,<sup>6</sup> in writing on the neurology of thinking, remarks that the transforms that govern the utilization of features of an excitation—what we speak of as "rules"—are determined by factors other than the nature of the immediately operative stimulus input. The question of what kinds of events bring about such alterations in rules for gating afferent stimulation is still an open one, and it is still a matter of debate whether there is some central activating system that is always and necessarily involved in such alterations.

#### A BRIEF CONCLUSION

The model of perception we have explored in these pages is one that is a drastic departure from the conventional stimulus-response, associational, or reflex-arc model that is a legacy to psy-

<sup>6</sup> See footnote 4.



chology from the neurophysiology of a past generation. It is a model in which, to use the language of Fessard (15), there is a de-emphasis on transmission of impulses across synaptic segments and a shift in emphasis to integration and autogenic activity—a model including complex networks with the capacity to hold up and to alter the characteristics of impulses transmitted to them, and with the capacity to initiate activity that is transmitted elsewhere to affect control of afferent impulses traveling to the cortex and efferent impulses traveling away from it. It is a system that, to put it figuratively, can determine within considerable limits the nature of the effective excitation which results when a change in physical energy impinges upon a sense receptor. The tracing and manipulation of efferent fibers carrying impulses to synapses along receptor pathways and to sensory receptors themselves indicate that the neural model we shall be using is one in which centrally induced control of sense data will play an increasing role. If the neurophysiology of a generation ago forced psychology into a peripheralist mold, certainly the model emerging today corrects this bias and places a new emphasis on the role of central factors in perception. Most important of all, I have the impression that the neurological model of perception that is now emerging begins for the first time to conform to our knowledge of more complex forms of perception, both in humans and at the infrahuman level.

## REFERENCES

1. ADRIAN, E. D. *The physical background of perception*. Oxford: Clarendon Press, 1947.
2. ADRIAN, E. D. The physiological basis of perception. In E. D. Adrian *et al.* (Eds.), *Brain mechanisms and consciousness*. Oxford: Blackwell, 1954.
3. ADRIAN, E. D., & MORUZZI, G. Cited in E. D. Adrian, *The physical background of perception*. Oxford: Clarendon Press, 1947.
4. BÉKÉSY, G. v. Description of some mechanical properties of the organ of Corti. *J. acoust. Soc. Amer.*, 1953, 25, 770-785.
5. BREMER, F. In E. D. Adrian *et al.* (Eds.), *Brain mechanisms and consciousness*. Oxford: Blackwell, 1954.
6. BRUNER, J. S., GOODNOW, J. J., & AUSTIN, G. A. *A study of thinking*. New York: Wiley, 1956.
7. BRUNER, J. S., & POSTMAN, L. On the perception of incongruity: a paradigm. *J. Pers.*, 1949, 18, 206-223.
8. CRAIK, K. *The nature of explanation*. Cambridge, England: Cambridge Univ. Press, 1952.
9. DAVIS, H. Psychophysiology of hearing and deafness. In S. S. Stevens (Ed.), *Handbook of experimental psychology*. New York: Wiley, 1951.
10. DETHIER, V. G. The dioptric apparatus of lateral ocelli. II. Visual capacities of the ocellus. *J. cell. comp. Physiol.*, 1943, 22, 115-126.
11. DITCHBURN, R. W. Report to the Experimental Psychology Group. Reading, England, January, 1957.
12. EXNER, S. *Die Physiologie der facettierten Augen von Krebsen und Insekten*. Leipzig und Wien, 1891.
13. FECHNER, G. T. *Elemente der Psychophysik*, I and II. Leipzig, 1869.
14. FESSARD, A. In E. D. Adrian *et al.* (Eds.), *Brain mechanisms and consciousness*. Oxford: Blackwell, 1954.
15. FESSARD, A. Mechanisms of nervous integration and conscious experience. In E. D. Adrian *et al.* (Eds.), *Brain mechanisms and consciousness*. Oxford: Blackwell, 1954.
16. FLETCHER, H. *Speech and hearing*. New York: D. Van Nostrand, 1929.
17. GALAMBOS, R. Suppression of auditory nerve activity by stimulation of efferent fibers to cochlea. *J. Neurophysiol.*, 1956, 19, 424-431.
18. GALAMBOS, R., & DAVIS, H. The response of single auditory-nerve fibers to acoustic stimulation. *J. Neurophysiol.*, 1943, 6, 39-58.
19. GALAMBOS, R., & DAVIS, H. Action potentials from single auditory nerve fibers? *Science*, 1948, 108, 513.
20. GALAMBOS, R., SHEATZ, G., & VERNIER, V. G. Electrophysiological correlates of a conditioned response in cats. *Science*, 1956, 123, 376-377.
21. GIBSON, J. J. *The perception of the visual world*. Boston: Houghton Mifflin, 1950.



22. GRANIT, R. *Receptors and sensory perception*. New Haven: Yale Univer. Press, 1955.
23. HEBB, D. O. *The organization of behavior*. New York: Wiley, 1949.
24. HEBB, D. O. The role of neurological ideas in psychology. *J. Pers.* 1951, 20, 40.
25. HEBB, D. O. The problem of consciousness and introspection. In E. D. Adrian *et al.* (Eds.), *Brain mechanisms and consciousness*. Oxford: Blackwell, 1954.
26. HELSON, H. Adaptation-level as a basis for a quantitative theory of frames of reference. *Psychol. Rev.*, 1948, 55, 297-313.
27. HERNANDEZ-PÉON, R., & SCHERRER, H. *Federation Proc.*, 1955, 14, 71.
28. HERNANDEZ-PÉON, R., SCHERRER, H., & JOUVET, M. Modification of electric activity in the cochlear nucleus during "attention" in unanesthetized cats. *Science*, 1956, 123, 331-332.
29. JAKOBSON, R., & HALLE, M. *Fundamentals of language*. Mouton: 'S-Gravenhagh (Holland), 1956.
30. KLÜVER, H. *Behavior mechanisms in monkeys*. Chicago: Univer. Chicago Press, 1933.
31. KUFFLER, S. W., & HUNT, C. C. The mammalian small nerve fibers: a system for efferent nervous regulation of muscle spindle discharge. *Proc. Assoc. Res. Nerv. Ment. Dis.*, 1952, Vol. 30.
32. LASHLEY, K. S. In search of the engram. *Symp. Soc. Exp. Biol.*, 1950, 4, 454-482.
33. LASHLEY, K. S. Dynamic processes in perception. In E. D. Adrian *et al.* (Eds.), *Brain mechanisms and consciousness*. Oxford: Blackwell, 1954.
34. LEKSELL, L. The action potential and excitatory effects of the small ventral root fibers to skeletal muscles. *Acta Physiol. Scand.*, 1945, 10, Suppl. 31.
35. MANDLER, G. Transfer of training as a function of degree of response overlearning. *J. exp. Psychol.*, 1954, 47, 411-417.
36. MCGILL, W. J., & ROSENBLITH, W. A. Electrical responses to two clicks: a simple statistical interpretation. *Bull. math. Biophysics*, 1951, 13, 69.
37. MILLER, G. A. Human memory and the storage of information. *Trans. on Information Theory, Inst. Radio Engrs.*, 1956, Vol. IT2, No. 3, 129-137.
38. PENFIELD, W. Studies of the cerebral cortex of man—a review and an interpretation. In E. D. Adrian *et al.* (Eds.), *Brain mechanisms and consciousness*. Oxford: Blackwell, 1954.
39. PFAFFMANN, C. Gustatory afferent impulses. *J. cell. comp. Physiol.*, 1941, 17, 243-258.
40. PIAGET, J. *Construction of reality in the child*. New York: Basic Books, 1954.
41. RIGGS, L. A. Visual acuity and the normal tremor of the eyes. *Science*, 1951, 114, 17-18.
42. ROSENBLITH, W. A. Electrical responses from the auditory nervous system. *Ann. Otology, Rhinology & Laryngology*, 1954, 63, 1-22.
43. SAUSSURE, F. DE. *Cours de linguistique general* (2nd ed.). Paris: Payot, 1922.
44. SENDEN, M. v. *Raum- und Gestaltaufassung bei operierten Blindgeborenen vor und nach der Operation*. Leipzig: Barth, 1932.
45. SHEARFLESS, S., & JASPER, H. Habituation of the arousal reaction. *Brain*, 1956, 79, 655-680.
46. STEVENS, S. S. On the psychophysical law. *Psychol. Rev.*, 1957, 64, 153-181.
47. TASAKI, I. Nerve impulses in the individual auditory nerve fibers of guinea pig. *J. Neurophysiol.*, 1954, 17, 97-122.
48. THORPE, W. H. *Learning and instinct in animals*. Cambridge: Harvard Univer. Press, 1956.
49. THOULESS, R. H. Individual differences in phenomenal regression. *Brit. J. Psychol.*, 1932, 22, 216-241.
50. TUNTURI, A. R. Audio frequency localization in the acoustic cortex of the dog. *Amer. J. Physiol.*, 1944, 141, 397-403.
51. TUNTURI, A. R. Further afferent connections to the acoustic cortex of the dog. *Amer. J. Physiol.*, 1945, 144, 389-394.
52. ULLYOT, P. The behaviour of *Dendrocoelum lacteum*. II. Responses in non-directional gradients. *J. exp. Biol.*, 1936, 13, 265-278.
53. WALLACH, H. Brightness constancy and the nature of achromatic colors. *J. exp. Psychol.*, 1948, 38, 310-324.
54. WEBER, E. H. In E. G. Boring, *Sensation and perception in the history of experimental psychology*. New York: D. Appleton-Century, 1942. Pp. 495 ff.
55. WOODWORTH, R. S. Reenforcement of perception. *Amer. J. Psychol.*, 1947, 60, 119-124.

(Received February 18, 1957)

## MOTIVATIONAL DETERMINANTS OF RISK- TAKING BEHAVIOR

JOHN W. ATKINSON<sup>1</sup>

*University of Michigan*

There are two problems of behavior which any theory of motivation must come to grips with. They may finally reduce to one; but it will simplify the exposition which follows to maintain the distinction in this paper. The first problem is to account for an individual's selection of one path of action among a set of possible alternatives. The second problem is to account for the amplitude or vigor of the action tendency once it is initiated, and for its tendency to persist for a time in a given direction. This paper will deal with these questions in a conceptual framework suggested by research which has used thematic apperception to assess individual differences in strength of achievement motivation (1, 14, 15).

The problem of selection arises in experiments which allow the individual to choose a task among alternatives that differ in difficulty (level of aspiration). The problem of accounting for the vigor of response arises in studies which seek to relate individual differences in strength of motivation to the level of performance when response output at a particular task is the dependent variable. In treating these two problems, the discussion will be constantly focused on the relationship of achievement motivation to risk-taking behavior, an important association uncovered by McClelland (14) in the investigation of the role of achievement motivation in entre-

preneurship and economic development.<sup>2</sup>

Earlier studies have searched for a theoretical principle which would explain the relationship of strength of motive, as inferred from thematic apperception, to overt goal-directed performance. The effect of situation cues (e.g., of particular instructions) on this relationship was detected quite early (1), and subsequent experiments have suggested a theoretical formulation similar to that presented by Tolman (21) and Rotter (20). It has been proposed that *n* Achievement scores are indices from thematic apperception are indices of individual differences in the strength of achievement motive, conceived as a relatively stable disposition to strive for achievement or success. This motive-disposition is presumed to be latent until aroused by situation cues which indicate that some performance will be instrumental to achievement. The strength of *aroused* motivation to achieve as manifested in performance has been viewed as a function of both the strength of motive and the *expectancy* of goal-attainment aroused by situation cues. This conception has provided a fairly adequate explanation of experimental results to date, and several of its implications have been tested (1, 2).

The similarity of this conception to the expectancy principle of performance developed by Tolman, which also takes account of the effects of a third variable, *incentive*, suggested the need for experiments to isolate the effects on motiva-

<sup>1</sup> I wish to acknowledge the stimulation and criticism of colleagues at the Center for Advanced Study in the Behavioral Sciences (1955-56), and also the current support for this research by a grant from the Ford Foundation.

<sup>2</sup> McClelland, D. C. Interest in risky occupations among subjects with high achievement motivation. Unpublished paper, Harvard University, June, 1956.



tion of variations in strength of expectancy of success and variations in the incentive value of particular accomplishments. The discussion which follows was prompted by the results of several exploratory experiments. It represents an attempt to state explicitly how individual differences in the strength of achievement-related motives influence behavior in competitive achievement situations. A theoretical model will be presented first, then a brief summary of some as yet unpublished experimental evidence will be introduced in order to call the reader's attention to the kinds of research problems it raises and the scope of its implications.

Three variables require definition and, ultimately, independent measurement. The three variables are *motive*, *expectancy*, and *incentive*. Two of these—expectancy and incentive—are similar to variables presented by Tolman (21) and Rotter (20). An expectancy is a cognitive anticipation, usually aroused by cues in a situation, that performance of some act will be followed by a particular consequence. The strength of an expectancy can be represented as the subjective probability of the consequence, given the act.

The incentive variable has been relatively ignored, or at best crudely defined, in most research. It represents the relative attractiveness of a specific goal that is offered in a situation, or the relative unattractiveness of an event that might occur as a consequence of some act. Incentives may be manipulated experimentally as, for example, when amount of food (reward) or amount of shock (punishment) is varied in research with animals.

The third variable in this triumvirate—motive—is here conceived differently than, for example, in the common conception of motivation as nondirective but energizing *drive* (3). A motive is conceived as a disposition to strive for

a certain kind of satisfaction, as a capacity for satisfaction in the attainment of a certain class of incentives. The names given motives—such as achievement, affiliation, power—are really names of classes of incentives which produce essentially the same kind of experience of satisfaction: pride in accomplishment, or the sense of belonging and being warmly received by others, or the feeling of being in control and influential. McClelland (13, pp. 341–352 and 441–458; 15) has presented arguments to support the conception of motives as relatively general and stable characteristics of the personality which have their origins in early childhood experience. The idea that a motive may be considered a *capacity for satisfaction* is suggested by Winterbottom's (15, 22) finding that children who are strong in achievement motive are rated by teachers as deriving more pleasure from success than children who are weak in achievement motive.

The general aim of one class of motives, usually referred to as appetites or approach tendencies, is to maximize satisfaction of some kind. The achievement motive is considered a disposition to approach success.

The aim of another class of motives is to minimize pain. These have been called aversions, or avoidant tendencies. An avoidance motive represents the individual's capacity to experience pain in connection with certain kinds of negative consequences of acts. The motive to avoid failure is considered a disposition to avoid failure and/or a capacity for experiencing shame and humiliation as a consequence of failure.

*The principle of motivation.* The strength of motivation to perform some act is assumed to be a multiplicative function of the strength of the motive, the expectancy (subjective probability) that the act will have as a consequence the attainment of an incentive, and the



value of the incentive:  $\text{Motivation} = f(\text{Motive} \times \text{Expectancy} \times \text{Incentive})$ . This formulation corresponds to Tolman's (21) analysis of performance except, perhaps, in the conception of a motive as a relatively stable disposition. When both motivation to approach and motivation to avoid are simultaneously aroused, the resultant motivation is the algebraic summation of approach and avoidance. The act which is performed among a set of alternatives is the act for which the resultant motivation is most positive. The magnitude of response and the persistence of behavior are functions of the strength of motivation to perform the act relative to the strength of motivation to perform competing acts.

Recent experiments (2) have helped to clarify one problem concerning the relationship between measures of the strength of a particular motive (*n* Achievement) and performance. Performance is positively related to the strength of a particular motive only when an expectancy of satisfying that motive through performance has been aroused, and when expectancies of satisfying other motives through the same action have not been sufficiently aroused to confound the simple relationship. This is to say no more than that, when expectancies of attaining several different kinds of incentives are equally *salient* in a situation, the determination of motivation to perform an act is very complex. Performance is then overdetermined in the sense that its strength is now a function of the several different kinds of motivation which have been aroused. The *ideal situation* for showing the relationship between the strength of a particular motive and behavior is one in which the only *reason* for acting is to satisfy that motive.

The theoretical formulation which follows pertains to such an *ideal achievement-related situation*, which is at best

only approximated in actual experimentation or in the normal course of everyday life. The discussion will deal only with the effects of the two motives, to achieve and to avoid failure, normally aroused whenever performance is likely to be evaluated against some standard of excellence.

*Behavior directed toward achievement and away from failure.* The problem of selection is confronted in the level-of-aspiration situation where the individual must choose among tasks which differ in degree of difficulty. The problem of accounting for the vigor of performance arises in the situation which will be referred to as *constrained performance*. Here there is no opportunity for the individual to choose his own task. He is simply given a task to perform. He must, of course, decide to perform the task rather than to leave the situation. There is a problem of selection. In referring to this situation as constrained performance, it is the writer's intention to deal only with those instances of behavior in which motivation for the alternative of leaving the situation is less positive or more negative than for performance of the task that is presented. Hence, the individual does perform the task that is given. The level of performance is the question of interest.

Elaboration of the implications of the multiplicative combination of motive, expectancy, and incentive, as proposed to account for strength of motivation, will be instructive if we can find some reasonable basis for assigning numbers to the different variables. The strength of expectancy can be represented as a subjective probability ranging from 0 to 1.00. But the problem of defining the positive incentive value of a particular accomplishment and the negative incentive value of a particular failure is a real stickler.

In past discussions of level of aspiration, Escalona and Festinger (see 12)

have assumed that, within limits, the attractiveness of success is a positive function of the difficulty of the task, and that the unattractiveness of failure is a negative function of difficulty, when the type of activity is held constant. The author will go a few steps farther with these ideas, and assume that degree of difficulty can be inferred from the subjective probability of success ( $P_s$ ). The task an individual finds difficult is one for which his subjective probability of success ( $P_s$ ) is very low. The task an individual finds easy is one for which his subjective probability of success ( $P_s$ ) is very high. Now we are in a position to make simple assumptions about the incentive values of success or failure at a particular task. Let us assume that the incentive value of success ( $I_s$ ) is a positive linear function of difficulty. If so, the value  $1 - P_s$  can represent  $I_s$ , the incentive value of success. When  $P_s$  is high (e.g., .90), an easy task,  $I_s$  is low (e.g., .10). When  $P_s$  is low (e.g., .10), a difficult task,  $I_s$  is high (e.g., .90). The negative incentive value of failure ( $I_f$ ) can be taken as  $-P_s$ . When  $P_s$  is high (e.g., .90), as in confronting a very easy task, the sense of humiliation accompanying failure is also very great (e.g.,  $-.90$ ).

However, when  $P_s$  is low (e.g., .10), as in confronting a very difficult task, there is little embarrassment in failing (e.g.,  $-.10$ ). We assume, in other words, that the (negative) incentive value of failure ( $I_f$ ) is a negative linear function of difficulty. It is of some importance to recognize the dependence of incentive values intrinsic to achievement and failure upon the subjective probability of success. One cannot anticipate the thrill of a great accomplishment if, as a matter of fact, one faces what seems a very easy task. Nor does an individual experience only a minor sense of pride after some extraordinary feat against what seemed to him overwhelming odds. The implications of the scheme which follows rest heavily upon the assumption of such a dependence.

In Table 1, values of 1 have been arbitrarily assigned to the achievement motive ( $M_s$ ) and the motive to avoid failure ( $M_f$ ). Table 1 contains the strength of motivation to approach success ( $M_s \times P_s \times I_s$ ) and motivation to avoid failure ( $M_f \times P_f \times I_f$ ) through performance of nine different tasks labeled A through I. The tasks differ in degree of difficulty as inferred from the subjective probability of success ( $P_s$ ). The incentive values of success and fail-

TABLE 1

AROUSED MOTIVATION TO ACHIEVE (APPROACH) AND TO AVOID FAILURE (AVOIDANCE) AS A JOINT FUNCTION OF MOTIVE ( $M$ ), EXPECTANCY ( $P$ ), AND INCENTIVE ( $I$ ), WHERE  $I_s = (1 - P_s)$  AND  $I_f = (-P_s)$

	Motivation to Achieve				Motivation to Avoid Failure				Resultant Motivation (Approach—Avoidance)
	$M_s \times P_s \times I_s = \text{Approach}$				$M_f \times P_f \times I_f = \text{Avoidance}$				
Task A	1	.10	.90	.09	1	.90	— .10	— .09	0
Task B	1	.20	.80	.16	1	.80	— .20	— .16	0
Task C	1	.30	.70	.21	1	.70	— .30	— .21	0
Task D	1	.40	.60	.24	1	.60	— .40	— .24	0
Task E	1	.50	.50	.25	1	.50	— .50	— .25	0
Task F	1	.60	.40	.24	1	.40	— .60	— .24	0
Task G	1	.70	.30	.21	1	.30	— .70	— .21	0
Task H	1	.80	.20	.16	1	.20	— .80	— .16	0
Task I	1	.90	.10	.09	1	.10	— .90	— .09	0



ure at each of the tasks have been calculated directly from the assumptions that incentive value of success equals  $1 - P_s$  and that incentive value of failure equals  $-P_s$ ; and  $P_s$  and  $P_f$  are assumed to add to 1.00.

Table 1 may be considered an extension of ideas presented in the *resultant valence* theory of level of aspiration by Escalona and Festinger (12). The present formulation goes beyond their earlier proposals (a) in making specific assumptions regarding the incentive values of success and failure, and (b) in stating explicitly how individual differences in strength of achievement motive and motive to avoid failure influence motivation.<sup>8</sup>

When the achievement motive is stronger ( $M_s > M_f$ ). The right-hand column of Table 1 shows the resultant motivation for each of the tasks in this special case where achievement motive and motive to avoid failure are equal in strength. In every case there is an approach-avoidance conflict with resultant motivation equal to 0. This means that if the achievement motive were stronger than the motive to avoid failure—for example, if we assigned  $M_s$  a value of 2—the resultant motivation would become positive for each of the tasks and its magnitude would be the same as in the column labeled *Approach*. Let us therefore consider only the strength of

approach motivation for each of the tasks, to see the implications of the model for the person in whom the need for achievement is stronger than his disposition to avoid failure.

One thing is immediately apparent. Motivation to achieve is strongest when uncertainty regarding the outcome is greatest, i.e., when  $P_s$  equals .50. If the individual were confronted with all of these tasks and were free to set his own goal, he should choose Task E where  $P_s$  is .50, for this is the point of maximum approach motivation. The strength of motivation to approach decreases as  $P_s$  increases from .50 to near certainty of success ( $P_s = .90$ ), and it also decreases as  $P_s$  decreases from .50 to near certainty of failure ( $P_s = .10$ ).

If this person were to be confronted with a single task in what is here called the constrained performance situation, we should expect him to manifest strongest motivation in the performance of a task of intermediate difficulty where  $P_s$  equals .50. If presented either more difficult tasks or easier tasks, the strength of motivation manifested in performance should be lower. The relationship between strength of motivation as expressed in performance level and expectancy of success at the task, in other words, should be described by a bell-shaped curve.

When the motive to avoid failure is stronger ( $M_f > M_s$ ). Let us now ignore the strength of approach motivation and tentatively assign it a value of 0, in order to examine the implications of the model for any case in which the motive to avoid failure is the stronger motive. The resultant motivation for each task would then correspond to the values listed in the column labeled *Avoidance*.

What should we expect of the person in whom the disposition to avoid failure is stronger than the motive to achieve? It is apparent at once that

<sup>8</sup> In the resultant valence theory of level of aspiration, the resultant force ( $f^*$ ) for a particular level of difficulty equals probability of success ( $P_s$ ) times valence of success ( $Va_s$ ) minus probability of failure ( $P_f$ ) times valence of failure ( $Va_f$ ). It is assumed that the valence of a goal [ $Va(G)$ ] depends partly on the properties of the activity and specific goal ( $G$ ) and partly on the state of need [ $t(G)$ ] of the person, [ $Va(G) = F(G, t(G))$ ] (11, p. 273). In the present conception, the relative rewarding or punishing properties of specific goals (i.e., incentives) and the more general disposition of the person toward a class of incentives (i.e., his motive) are given independent status.



the resultant motivation for every task would be negative for him. This person should want to avoid all of the tasks. Competitive achievement situations are unattractive to him. If, however, he is constrained (e.g., by social pressures) and asked to set his level of aspiration, he should *avoid* tasks of intermediate difficulty ( $P_s = .50$ ) where the arousal of anxiety about failure is greatest. He should choose either the easiest ( $P_s = .90$ ) or the most difficult task ( $P_s = .10$ ). The strength of avoidant motivation is weakest at these two points.

In summary, the person in whom the achievement motive is stronger should set his level of aspiration in the intermediate zone where there is moderate risk. To the extent that he has any motive to avoid failure, this means that he will voluntarily choose activities that *maximize* his own anxiety about failure! On the other hand, the person in whom the motive to avoid failure is stronger should select either the easiest of the alternatives or should be extremely speculative and set his goal where there is virtually no chance for success. These are activities which *minimize* his anxiety about failure.

How does the more fearful person behave when offered only a specific task to perform? He can either perform the task or leave the field. If he chooses to leave the field, there is no problem. But if he is constrained, as he must be to remain in any competitive achievement situation, he will stay at the task and presumably work at it. But how hard will he work at it? He is motivated to avoid failure, and when constrained, there is only one path open to him to avoid failure—success at the task he is presented. So we expect him to manifest the strength of his motivation to avoid failure in performance of the task. He, too, in other words, should

*try hardest*<sup>4</sup> when  $P_s$  is .50 and less hard when the chance of winning is either greater or less. The 50–50 alternative is the last he would choose if allowed to set his own goal, but once constrained he must try hard to avoid the failure which threatens him. Not working at all will guarantee failure of the task. Hence, the thought of not working at all should produce even stronger avoidant motivation than that aroused by the task itself.

In other words, irrespective of whether the stronger motive is to achieve or to avoid failure, the strength of motivation to perform a task when no alternatives are offered and when the individual is constrained should be greatest when  $P_s$  is .50. This is the condition of greatest uncertainty regarding the outcome. But when there are alternatives which differ in difficulty, the choice of level of aspiration by persons more disposed to avoid failure is diametrically opposite to that of persons more disposed to seek success. The person more motivated to achieve should prefer a moderate risk. His level of aspiration will fall at the point where his positive motivation is strongest, at the point where the odds seem to be 50–50. The fearful person, on the other hand, must select a task even though all the alternatives are threatening to him. He prefers the least threatening of the available alternatives: either the task which is so easy he cannot fail, or the task which is so difficult that failure would be no cause for self-blame and embarrassment.

The tendency for anxious persons to set either extremely high or very low aspirations has been noted over and over

<sup>4</sup> I do not mean to exclude the possibility that the very anxious person may suffer a performance decrement due to the arousal of some "task-irrelevant" avoidant responses, as proposed in the interpretation of research which has employed the Mandler-Sarason Measure of Test Anxiety (16).

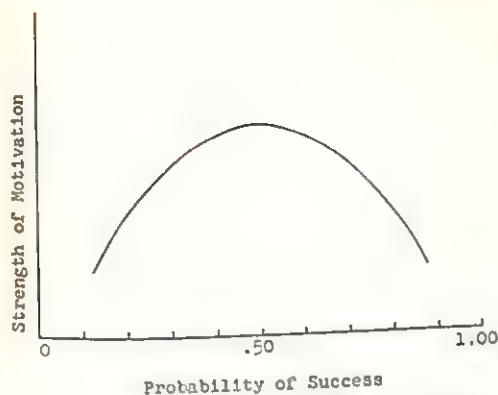


FIG. 1. Strength of motivation to achieve or to avoid failure as a function of the subjective probability of success, i.e., the difficulty of the task.

again in the literature on level of aspiration (12). Typically, groups of persons for whom the inference of greater anxiety about failure seems justified on the basis of some personality assessment show a much greater variance in level of aspiration than persons whose motivation is inferred to be more normal or less anxious. When the details of behavior are examined, it turns out that they are setting their aspiration level either *defensively* high or *defensively* low.

Without further assumptions, the theory of motivation which has been presented when applied to competitive-achievement activity implies that the relationship of constrained performance to expectancy of goal-attainment should take the bell-shaped form shown in Fig. 1, whether the predominant motive is to achieve or to avoid failure. Further, the theory leads to the prediction of exactly opposite patterns for setting the level of aspiration when the pre-dominant motivation is approach and when it is avoidant, as shown in Fig. 2.

Both of these hypotheses have been supported in recent experiments. The writer<sup>5</sup> offered female college students

a modest monetary prize for good performance at two 20-minute tasks. The probability of success was varied by instructions which informed the subject of the number of persons with whom she was in competition and the number of monetary prizes to be given. The stated probabilities were  $\frac{1}{20}$ ,  $\frac{1}{8}$ ,  $\frac{1}{2}$ , and  $\frac{3}{4}$ . The level of performance was higher at the intermediate probabilities than at the extremes for subjects having high thematic apperceptive *n* Achievement scores, and also for subjects who had low *n* Achievement scores, presumably a more fearful group.

McClelland<sup>6</sup> has shown the diametrically opposite tendencies in choice of level of aspiration in studies of children in kindergarten and in the third grade. One of the original level-of-aspiration experiments, the ring-toss experiment, was repeated with five-year-olds, and a non-verbal index of the strength of achieve-

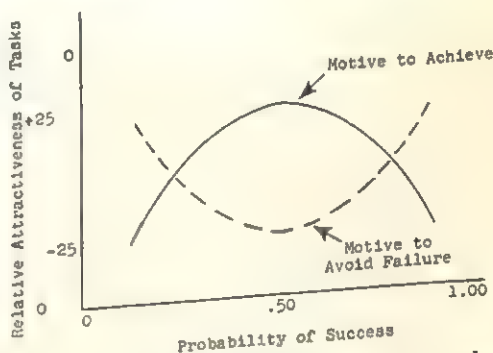


FIG. 2. Relative attractiveness of tasks which differ in subjective probability of success (i.e., in difficulty). The avoidance curve has been inverted to show that very difficult and very easy tasks arouse less fear of failure and hence are less unattractive than moderately difficult tasks.

ment motives, expectancies, and incentives. To appear in *Motives in fantasy, action, and society*. Princeton: Van Nostrand (in preparation).

<sup>6</sup> McClelland, D. C. Risk-taking in children with high and low need for achievement. To appear in *Motives in fantasy, action, and society*. Princeton: Van Nostrand (in preparation).

<sup>5</sup> Atkinson, J. W. Towards experimental analysis of human motivation in terms of



ment motive was employed. Children who were high in  $n$  Achievement more frequently set their level of aspiration in the intermediate range of difficulty. They took more shots from a modest distance. Children who were low in  $n$  Achievement showed a greater preponderance of choices at the extreme levels of difficulty. They more often stood right on top of the peg or stood so far away that success was virtually impossible. The same difference between high and low  $n$  Achievement groups was observed on another task with children in the third grade. McClelland views these results as consistent with his theoretical argument concerning the role of achievement motivation in entrepreneurship and economic development (14). He has called attention to the relationship between achievement motivation and an interest in enterprise which requires moderate or calculated risks, rather than very safe or highly speculative undertakings.

In an experiment designed for another purpose, Clark, Teevan, and Ricciuti (4) have presented results with college students comparable to those of McClelland. Immediately before a final examination in a college course, students were asked a series of questions pertaining to grade expectations, affective reactions to grades, and the grades they would *settle for* if excused from taking the exam. A number of indices were derived from responses to these questions, by which the students were classified as: *hopeful of success*, i.e., if the *settle-for* grade was near the maximum grade the student thought he could possibly achieve; *fearful of failure*, i.e., if the *settle-for* grade was near the minimum grade the student thought he might possibly drop to; and *intermediate*, i.e., if the *settle-for* grade fell somewhere between these two extremes. Previously obtained  $n$  Achievement scores were significantly higher for

the *intermediate* group than for the two groups who set either extremely high or low levels of aspiration.

In terms of the model presented in Table 1, the two extreme patterns of aspirant behavior which are here designated *hope of success* and *fear of failure* are to be considered two *phenotypically* dissimilar alternatives that are *genotypically* similar. That is, they both function to avoid or reduce anxiety for the person in whom the motive to avoid failure is stronger than the motive to achieve.

A question may arise concerning the legitimacy of inferring relatively stronger motive to avoid failure from a low  $n$  Achievement score in thematic apperception. The inference seems justified on several counts. First, the kind of learning experience which is thought to contribute to the development of a positive motive to achieve (15, 22) seems incompatible with the kind of experience which would contribute to the development of an avoidant motive. In any specific early learning experience in which successful independent accomplishment is encouraged and rewarded, it seems impossible for incompetence, at the same time, to be punished. Second, even if it is assumed that high and low  $n$  Achievement groups may be equal in the disposition to be fearful of failure, the fact that one group does not show evidence of a strong motive to achieve (the group with low  $n$  Achievement scores) suggests that fear of failure should be *relatively* stronger in that group than in the group which does show evidence of strong  $n$  Achievement (high  $n$  Achievement scores). Finally, Raphelson (19) has presented evidence that  $n$  Achievement, as measured in thematic apperception, is *negatively* related to both scores on the Mandler-Sarason Scale of Test Anxiety and a psychogalvanic index of manifest anxiety obtained in a test situation. Test



anxiety scores and the psychogalvanic index of manifest anxiety were *positively* correlated, as they should be if each is an effective measure of fear aroused in a competitive situation.

Although a low  $n$  Achievement score can hardly be viewed as a direct index of the disposition to avoid failure, there seems good presumptive evidence that fear of failure is *relatively* stronger than the achievement motive in such a group. And this presumption is all the theory demands to explain the pattern of goal setting which focuses upon the extremes in the range of difficulty among persons low in  $n$  Achievement.

The details of the exploratory experiments suggest that one further assumption be made. In both experiments, the high  $n$  Achievement groups showed evidence of maximum motivation when the observed or stated probability of success was approximately .33. At this point, the high  $n$  Achievement group showed the highest level of constrained performance. And this point was most favored by the high  $n$  Achievement group in setting level of aspiration in the McClelland experiment. The assumption to be made seems a reasonable one: the relative strength of a motive influences the subjective probability of the consequence consistent with that motive—i.e., biases it upwards. In other words, the stronger the achievement motive relative to the motive to avoid failure, the higher the subjective probability of success, given stated odds. The stronger the motive to avoid failure relative to the achievement motive, the higher the subjective probability of failure, given stated odds or any other objective basis for inferring the strength of expectancy. Some evidence from two earlier studies is pertinent. When subjects stated the score that they *expected* to make on a test with very ambiguous or conflicting cues from past performance (15, p. 247) or when faced with a novel task at

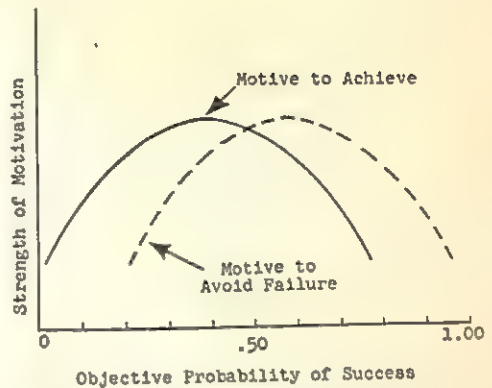


FIG. 3. Strength of motivation to achieve and to avoid failure as a function of the *objective* probability of success. It is assumed that the subjective probability of the consequence consistent with the stronger motive is biased upwards.

which they had no experience (18), the stated level of *expectation* was positively related to  $n$  Achievement. The biasing effect of the motive on subjective probability should diminish with repeated learning experience in the specific situation.

When this assumption is made, the point of maximum motivation to achieve now occurs where the stated (objective) odds are somewhat *lower* than .50; and the point of maximum motivation to avoid failure occurs at a point somewhat higher than stated odds of .50, as shown in Fig. 3. The implications of this assumption for constrained performance in somewhat novel situations are evident in the figure. When the achievement motive is stronger than the motive to avoid failure, there should be a tendency for stronger motivation to be expressed in performance when the objective odds are long, i.e., below .50. When the motive to avoid failure is stronger than the achievement motive, there should be greater motivation expressed when the objective odds are short, i.e., above .50.

*The effects of success and failure.* Let us return to the model and ask, What are the effects of success and fail-

ure on the level of motivation? We may refer back to Table 1 to answer this question. First, let us consider the effects of success or failure on the level of motivation in a person whose motive to achieve is stronger than his motive to avoid failure. In the usual level-of-aspiration situation, he should initially set his goal where  $P_s$  equals .50. In Table 1, this is Task E. If he succeeds at the task,  $P_s$  should increase. And, assuming that the effects of success and failure generalize to similar tasks, the  $P_s$  at Task D which was initially .40 should increase toward .50. On the next trial,  $P_s$  at Task E is now greater than .50, and  $P_s$  at Task D now approaches .50. The result of this change in  $P_s$  is diminished motivation to achieve at the old task, E, and increased motivation to achieve at Task D, *an objectively more difficult task*. The observed level of aspiration should increase in a step-like manner following success, because there has been a change in motivation.

A further implication of the change in strength of motivation produced by the experience of success is of great consequence: given a single, very difficult task (e.g.,  $P_s = .10$ ), the effect of continued success in repeated trials is first a gradual increase in motivation as  $P_s$  increases to .50, followed by a gradual decrease in motivation as  $P_s$  increases further to the point of certainty ( $P_s = 1.00$ ). Ultimately, as  $P_s$  approaches 1.00, satiation or loss of interest should occur. The task no longer arouses any motivation at all. Why? Because the subjective probability of success is so high that the incentive value is virtually zero. Here is the clue to understanding how the achievement motive can remain insatiable while satiation can occur for a particular line of activity. The strength of motive can remain unchanged, but interest in a particular task can diminish com-

pletely. Hence, when free to choose, the person who is stronger in achievement motive should always look for new and more difficult tasks as he masters old problems. If constrained, the person with a strong achievement motive should experience a gradual loss of interest in his work. If the task is of intermediate difficulty to start with ( $P_s = .50$ ), or is definitely easy ( $P_s > .50$ ), his interest should begin to wane after the initial experience of success.

But what of the effect of failure on the person who is more highly motivated to achieve than to avoid failure? Once more we look at the *Approach* column of Table 1. If he has chosen Task E ( $P_s = .50$ ) to start with and fails at it, the  $P_s$  is reduced. Continued failure will mean that soon Task F (formerly  $P_s = .60$ ) will have a  $P_s$  near .50. He should shift his interest to this task, which was *objectively less difficult* in the initial ordering of tasks. This constitutes what has been called a lowering of the level of aspiration. He has moved to the easier task as a consequence of failure.

What is the effect of continued failure at a single task? If the initial task is one that appeared relatively easy to the subject (e.g.,  $P_s = .80$ ) and he fails, his motivation should increase! The  $P_s$  will drop toward .70, but the incentive value or attractiveness of the task will increase. Another failure should increase his motivation even more. This will continue until the  $P_s$  has dropped to .50. Further failure should then lead to a gradual weakening of motivation as  $P_s$  decreases further. In other words, the tendency of persons who are relatively strong in achievement motive to persist at a task in the face of failure is probably attributable to the relatively high subjective probability of success, initially. Hence, failure has the effect of increasing the strength of their motivation, at least for a time. Ultimately,



however, interest in the task will diminish if there is continued failure. If the initial task is perceived by the person as very difficult to start with ( $P_s < .50$ ), motivation should begin to diminish with the first failure.

Let us turn now to the effect of success and failure on the motivation of the person who is more strongly disposed to be fearful of failure. If the person in whom the motive to avoid failure is stronger has chosen a very difficult task in setting his level of aspiration (e.g., Task A where  $P_s = .10$ ) and succeeds,  $P_s$  increases and his motivation to avoid the task is paradoxically increased! It would almost make sense for him deliberately to fail, in order to keep from being faced with a stronger threat on the second trial. If there are more difficult alternatives, he should raise his level of aspiration to avoid anxiety! Fortunately for this person, his strategy (determined by the nature of his motivation) in choosing a very difficult task to start with protects him from this possibility, because  $P_s$  is so small that he will seldom face the paradoxical problem just described. If he fails at the most difficult task, as is likely,  $P_s$  decreases further,  $P_f$  increases further, and the aroused motivation to avoid failure is reduced. By continued failure he further reduces the amount of anxiety about failure that is aroused by this most difficult task. Hence, he should continue to set his level at this point. If he plays the game long enough and fails continuously, the probability of failure increases for all levels of difficulty. Sooner or later the minimal motivation to avoid failure at the most difficult task may be indistinguishable from the motivation to avoid failure at the next most difficult task. This may ultimately allow him to change his level of aspiration to a somewhat less difficult task without acting in gross contra-

diction to the proposed principle of motivation.

If our fearful subject has initially chosen the easiest task (Task I where  $P_s = .90$ ) and if he fails,  $P_s$  decreases toward .80, and his motivation to avoid the task also increases. If there is no easier task, the most difficult task should now appear least *unattractive* to him, and he should jump from the easiest to the most difficult task. In other words, continued failure at a very easy task decreases  $P_s$  toward .50; and, as Table 1 shows, a change of this sort is accompanied by increased arousal of avoidant motivation. A wild and apparently irrational jump in level of aspiration from very easy to very difficult tasks, as a consequence of failure, might be mistakenly interpreted as a possible effort on the part of the subject to gain social approval by seeming to set high goals. The present model predicts this kind of activity without appealing to some extrinsic motive. It is part of the strategy of minimizing expected pain of failure after one has failed at the easiest task.

If our fear-disposed subject is successful at the most simple task, his  $P_s$  increases, his  $P_f$  decreases, and his motivation to avoid this task decreases. The task becomes less and less unpleasant. He should continue playing the game with less anxiety.

Table 1, when taken in its entirety, deals with the special case of the person in whom the two motives are exactly equal in strength. The implications are clear. In the constrained-performance situation, he should work hardest when the probability of success is .50, because motivation to achieve and motivation to avoid failure will summate in the constrained instrumental act which is at the same time the pathway toward success and away from failure. (This summation should also occur in the cases where one motive is stronger.) But in the level-of-aspiration setting



where there is an opportunity for choice among alternatives, the avoidance motivation exactly cancels out the approach motivation. Hence, the resultant motivation for each of the alternatives is zero. His choice of level of aspiration cannot be predicted from variables intrinsic to the achievement-related nature of the task. If there is any orderly pattern in this conflicted person's level of aspiration, the explanation of it must be sought in extrinsic factors, e.g. *the desire to gain social approval*. Such a desire can also be conceptualized in terms of motive, expectancy, and incentive, and the total motivation for a particular task can then be attributed to both achievement-related motives and other kinds of motives engaged by the particular features of the situation.

In recent years there has been something of a rebirth of interest in the problems of level of aspiration, particularly in pathological groups. The tendency for anxious groups to show much greater variability in level of aspiration, setting their goals either very high or very low relative to less anxious persons, was noted in early studies by Sears, Rotter, and others (12). Miller (17), Himmelweit (9), and Eysenck and Himmelweit (8) have produced substantial evidence that persons with affective disorders (neurasthenia or dysthymia) typically set extremely high goals for themselves; hysterics, on the other hand, show a minimal level of aspiration, often setting their future goal even below the level of past performance. In all of these studies, normal control groups have fallen between these two extremes, as might be expected from the present model if *normals* are relatively more positive in their motivation in achievement-related situations.

In the work of Eysenck (7) and his colleagues, both dysthymics and hysterics show greater *neuroticism* than nor-

mal subjects. Eysenck's interpretation of this factor as autonomic sensitivity is consistent with the implications of the present model, which attributes the setting of extremely high or low levels of aspiration to relatively strong motivation to avoid failure. A second factor, *extraversion-introversion*, discriminates the affective disorders and hysterics where the present model, dealing only with motives intrinsic to the competitive achievement situation, does not. An appeal to some other motivational difference, e.g., in strength of *n* Affiliation, might also predict the difference in pattern of level of aspiration.

### *Probability Preferences*

The present analysis is relevant to another domain of current research interest, that specifically concerned with the measurement of subjective probability and utility. Edwards (5, 6), for example, has reported probability preferences among subjects offered alternative bets having the same expected value. We<sup>7</sup> have repeated the Edwards type experiment (e.g., 6/6 of winning 30¢ versus 1/6 of winning \$1.80) with subjects having high and low *n* Achievement scores. The results show that persons high in *n* Achievement more often prefer intermediate probabilities (4/6, 3/6, 2/6) to extreme probabilities (6/6, 5/6, 1/6) than do persons low in *n* Achievement. What is more, the same differential preference for intermediate risk was shown by these *same* subjects when they were allowed to choose the distance from the target for their shots in a shuffleboard game. In other words, the incentive values of winning qua winning, and losing qua losing, presumably developed in achievement activities early in life, generalize to the gambling

<sup>7</sup> Atkinson, J. W., Bastian, J. R., Earl, R. W., and Litwin, G. H. The achievement motive, goal-setting, and probability preferences (in preparation).

situation in which winning is really *not* contingent upon one's own skill and competence.

### *Social Mobility Aspirations*

Finally, the present model may illuminate a number of interesting research possibilities having to do with social and occupational mobility. The ranking of occupations according to their prestige in Western societies clearly suggests that occupations accorded greater prestige are also more difficult to attain. A serious effort to measure the perceived probability of being able to attain certain levels on the occupational ladder should produce a high negative correlation with the usual ranking on prestige. If so, then the present model for level of aspiration, as well as its implications for persons who differ in achievement-related motives, can be applied to many of the sociological problems of mobility aspirations. A recent paper by Hyman (10) has laid the groundwork for such an analysis.

### SUMMARY

A theoretical model is presented to explain how the motive to achieve and the motive to avoid failure influence behavior in any situation where performance is evaluated against some standard of excellence. A conception of motivation in which strength of motivation is a joint multiplicative function of motive, expectancy (subjective probability), and incentive is offered to account for the selection of one task among alternatives which differ in difficulty (level of aspiration), and also to account for performance level when only one task is presented. It is assumed that the incentive value of success is a positive linear function of difficulty as inferred from the subjective probability of success; and negative incentive value of failure is assumed to be a negative linear function of difficulty. The major im-

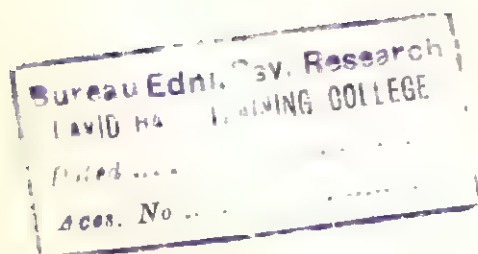
plications of the theory are (a) that performance level should be greatest when there is greatest uncertainty about the outcome, i.e., when subjective probability of success is .50, whether the motive to achieve or the motive to avoid failure is stronger within an individual; but (b) that persons in whom the achievement motive is stronger should prefer intermediate risk, while persons in whom the motive to avoid failure is stronger should avoid intermediate risk, preferring instead either very easy and safe undertakings or extremely difficult and speculative undertakings. Results of several experiments are cited, and the implications of the theoretical model for research on probability preferences in gambling and studies of social mobility aspirations are briefly discussed.

### REFERENCES

1. ATKINSON, J. W. Explorations using imaginative thought to assess the strength of human motives. In M. R. Jones (Ed.), *Nebraska Symposium on Motivation*, 1954. Lincoln, Nebr.: Univ. of Nebraska Press, 1954.
2. ATKINSON, J. W., & REITMAN, W. R. Performance as a function of motive strength and expectancy of goal-attainment. *J. abnorm. soc. Psychol.*, 1956, 53, 361-366.
3. BROWN, J. S. Problems presented by the concept of acquired drives. In *Current theory and research in motivation*. Lincoln, Nebr.: Univ. of Nebraska Press, 1953.
4. CLARK, R. A., TEEVAN, R., & RICCIUTI, H. N. Hope of success and fear of failure as aspects of need for achievement. *J. abnorm. soc. Psychol.*, 1956, 53, 182-186.
5. EDWARDS, W. Probability preferences in gambling. *Amer. J. Psychol.*, 1953, 66, 349-364.
6. EDWARDS, W. The theory of decision making. *Psychol. Bull.*, 1954, 51, 380-417.
7. EYSENCK, H. J. A dynamic theory of anxiety and hysteria. *J. ment. Sci.*, 1955, 101, 28-51.
8. EYSENCK, H. J., & HIMMELWEIT, H. T. An experimental study of the reactions

- of neurotics to experiences of success and failure. *J. gen. Psychol.*, 1946, 35, 59-75.
9. HIMMELWEIT, H. T. A comparative study of the level of aspiration of normal and neurotic persons. *Brit. J. Psychol.*, 1947, 37, 41-59.
  10. HYMAN, H. H. The value systems of different classes: a social psychological contribution to the analysis of stratification. In R. Bendix & S. M. Lipset (Eds.), *Class, status, and power*. Glencoe, Ill.: Free Press, 1953.
  11. LEWIN, K. *Field theory in social science*. D. Cartwright (Ed.). New York: Harper Bros., 1951.
  12. LEWIN, K., DEMBO, T., FESTINGER, L., & SEARS, P. S. Level of aspiration. In J. McV. Hunt (Ed.), *Personality and the behavior disorders*, Vol. 1, Chap. 10. New York: Ronald Press, 1944.
  13. MCCLELLAND, D. C. *Personality*. New York: William Sloane Associates, 1951.
  14. MCCLELLAND, D. C. Some social consequences of achievement motivation. In M. R. Jones (Ed.), *Nebraska symposium on motivation*, 1955. Lincoln: Univer. of Nebraska Press, 1955.
  15. MCCLELLAND, D. C., ATKINSON, J. W., CLARK, R. A., & LOWELL, E. L. *The achievement motive*. New York: Appleton-Century-Crofts, 1953.
  16. MANDLER, G., & SARASON, S. B. A study of anxiety and learning. *J. abnorm. soc. Psychol.*, 1952, 16, 115-118.
  17. MILLER, D. R. Responses of psychiatric patients to threat of failure. *J. abnorm. soc. Psychol.*, 1951, 46, 378-387.
  18. POTTHARST, B. C. The achievement motive and level of aspiration after experimentally induced success and failure. Unpublished doctor's dissertation, Univer. of Michigan, 1956.
  19. RAPHELSON, A. Imaginative and direct verbal measures of anxiety related to physiological reactions in the competitive achievement situation. Unpublished doctor's dissertation, Univer. of Michigan, 1956.
  20. ROTTER, J. B. *Social learning and clinical psychology*. New York: Prentice-Hall, 1954.
  21. TOLMAN, E. C. Principles of performance. *Psychol. Rev.*, 1955, 62, 315-326.
  22. WINTERBOTTOM, M. R. The relation of childhood training in independence to achievement motivation. Unpublished doctor's dissertation, Univer. of Michigan, 1952.

(Received February 7, 1957)





## DURABLE SECONDARY REINFORCEMENT: METHOD AND THEORY<sup>1</sup>

DONALD W. ZIMMERMAN

*University of Illinois*

For at least a decade now, the importance of secondary reinforcement has been increasingly emphasized by learning theorists and by systematists in general. It is true that Spence (22) has recently revised his earlier position (21) in which this phenomenon played an important role. And McClelland (11) has for some time questioned its importance because of its apparent instability. Nevertheless, secondary reinforcement is alluded to more and more frequently in the psychological literature; and an empirical demonstration of its reality and durability, if reality and durability it has, becomes correspondingly more desirable.

The purpose of the present paper is to report a method whereby secondary reinforcement can be made extraordinarily effective and stable. Some of the theoretical implications of this finding will also be briefly discussed.

### I. A TWO-STAGE INTERMITTENT RE- INFORCEMENT PROCEDURE WHICH LEADS TO POWERFUL SECONDARY REINFORCEMENT

With the experimental procedure to be reported here, it has been found possible to give a secondary reinforcer such strength that it could, if desired, be used as a class demonstration of the Law of Effect, much in the manner of such primary reinforcers as food and water. The method whereby such stable sec-

ondary reinforcement can be established is as follows:

*Training.* A thirsty rat is put into a small (11 × 11 × 11-in.) box fitted with a motor-driven water delivery mechanism. After the animal becomes habituated to the situation, a formerly neutral stimulus is given secondary reinforcement potential in this way: A two-second buzzer, followed by operation of the water delivery dipper, is presented at approximately one-minute intervals. Buzzer presentations are programmed independently of the behavior of the animal, with the exception that the buzzer never sounds while the animal hovers over the water-delivery aperture.

After the approach response has been firmly established, water reinforcements are omitted following the buzzer, at first on alternate presentations, then successively in longer runs. The ratio of rewarded to nonrewarded presentations varies somewhat randomly, the mean number of reinforcement omissions, however, moving upward. A 1:10 ratio is finally stabilized, with the longest single nonreinforced run being 1:14. What is involved here, then, is a *variable ratio* of water-rewarded to nonrewarded S<sup>r</sup> (secondary-reinforcement stimulus) presentations,<sup>2</sup> this ratio being gradually increased. During this training the vigor with which the animal leaps to the water delivery aperture following the buzzer does not lessen, in spite of the fact that the frequency of "pay off" is progressively decreasing.

<sup>1</sup> The author wishes to express his gratitude to Dr. O. H. Mowrer for offering many helpful suggestions during the course of the research here reported and for reading and criticizing the manuscript.

<sup>2</sup> For convenience the following symbols will be used: S<sup>r</sup> for secondary reinforcer, S<sup>D</sup> for positive discriminative stimulus, S<sup>A</sup> for negative discriminative stimulus.

During this type of training what happens is that the buzzer becomes a signal "that water *may* now be available." Water is received at no time *except when the buzzer sounds*, so the animal soon learns to "pay attention" to the buzzer—as it were, to "jump at the chance" which it affords. But the buzzer is not a sure sign of water.

**Testing.** Following the training procedure just described, the buzzer is tested for its secondary reinforcing properties by presenting it as a consequence of the rat's pressing a small bar which is now available. The buzzer ( $S^r$ ) alone is thus used to strengthen this response, no further water reinforcements being given. However, instead of having the buzzer invariably follow bar pressing, the procedure again involves intermittent reinforcement. Here various options regarding the detailed procedure that may be used are open to the experimenter. The one that produced the results shown in Fig. 1 went as follows. On the first day of testing, the bar-pressing behavior was established by reinforcing (with the buzzer) the first six responses (segment A). Thereafter  $S^r$  was given following a response only when at least a minute had elapsed since the preceding reinforcement (segment B). When the buzzer was discontinued entirely, a typical extinction performance followed (segment C). Reconditioning, with secondary reinforcement only, was possible on a second day, without intervening primary reinforcement of the secondary reinforcer. Various fixed-ratio schedules have also given satisfactory results.

A point to be stressed here is that the intermittent reinforcement schedules of the kind just suggested for use during the test period will increase response output *provided that, on the basis of the prior training procedure (involving intermittent primary reinforcement), the secondary reinforcer itself has been*

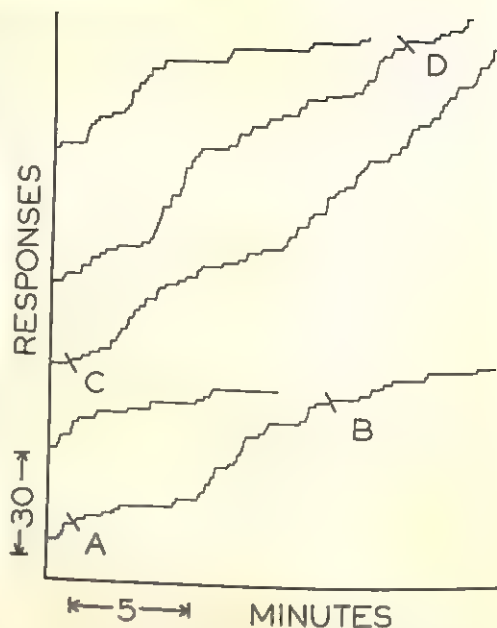


FIG. 1. Cumulative curves for acquisition and extinction of a bar-pressing response which received only secondary reinforcement. The two lower curves are for the animal's performance on Day 1; the three upper curves are for Day 2. From the origin to A, all responses were reinforced. From A to B, a one-minute fixed-interval schedule was used. Extinction began at B. At C (second day), the fixed-interval reinforcement schedule was resumed. Extinction began again at D. Each step rise in the curves represents one bar response.

*made "strong enough" to withstand a schedule of this kind.* If  $S^r$  is relatively weak, as in the usual sort of secondary reinforcement experiment, its effectiveness will wear out before any such schedule has a chance to operate.

Thus, two kinds of intermittent reinforcement are involved in the experimental procedure just described. One is intermittent reinforcement in the establishment of the secondary reinforcer, i.e., intermittent association of primary reinforcement with the secondary reinforcer. The other is intermittent administration of the secondary reinforcer itself, following the occurrence of the new response to be thereby strengthened. This second intermittent pro-

cedure may follow any of various schedules, typical results for one of which are shown.

The distinctive characteristics of the foregoing experimental procedure may be listed as follows:

1. The secondary reinforcing stimulus,  $S^r$ , is *discrete*. That is, it is not, as in many studies, some fixed characteristic of the experimental situation which persists throughout time—such as a continuous light in a Skinner box (4, 5), or the color of a runway goal box (16). In this respect it parallels a typical primary reinforcing event, such as the appearance (and disappearance) of a bit of food or water. The click of a food-delivery magazine, for a hungry rat, is a discrete event which may be presented anew following selected instances of a new response which is to be strengthened. Moreover, the stimulus thus used has an independently observable effect on behavior *before* the test procedure begins. The sound of magazine operation, following bar pressing (in an ordinary Skinner box), releases food-approaching behavior. In the same way the stimulus (buzzer) used in the present experiment leads to water-aperture approach during the training period, before its secondary reinforcing properties are independently demonstrated. In the case of something such as the color of a goal box, no such initial control of behavior is observed.

2. During training, water never appears without being preceded by  $S^r$ , but the appearance of  $S^r$  does not insure the appearance of water. In this respect the procedure is just the opposite of that used by Melching (12), where secondary reinforcement implies primary reinforcement but not the reverse.

3. The intermittent-reinforcement ratios used (both in the association of  $S^r$  with water and the presentation of  $S^r$  following the bar response) are relatively high. Most experimenters have

used ratios of the order of 1:2 or 1:3. In this respect the present procedure differs from that used by Saltzman (16) and Clayton (2), but is like that of Dinsmoor (4, 5). In order to achieve a high ratio of reinforcement, approximation by successive steps appears to be essential.

4. The reinforcing effect of  $S^r$  is clearly exhibited in the behavior of an *individual animal*. The observed experimental effect is so large that it is not necessary to use statistical criteria to demonstrate its reality. The same clear-cut effect has been obtained over and over again, with many animals.

## II. HOW DOES INTERMITTENT REINFORCEMENT INCREASE RESISTANCE TO EXTINCTION?

In the preceding section a method, admittedly rather complex, has been reported whereby highly stable and powerful secondary reinforcement can be demonstrated. This method is successful because it makes the phenomenon in question unusually resistant to extinction, and this end is achieved by the use of schedules of partial reinforcement. In order to get a better theoretical grasp of what is involved here, it will be useful to review a part of the history of the partial-reinforcement procedure.

The notion that failure or adversity, if ultimately followed by success, strengthens an individual's "character" and capacity to persevere is a very old one. The writer of the Book of James in the New Testament shows a clear appreciation of it. And Jenkins and Stanley, in their 1950 review, have this to say:

Partial reinforcement is such a ubiquitous phenomenon that one wonders why psychologists for so long failed to examine its consequences. . . . The animal trainer feeds his dogs only after several occurrences of the desired act, the clerk receives his check only once each month, the psychologist publishes his



findings only after prolonged endeavor, the gambler continues to bet despite an infrequent payoff, *ad infinitum*. Continuous reinforcement seems to be the exception, not the rule (10, p. 196).

When Skinner (19), first, and then Humphreys (9) and others reported experimental evidence on this score, their findings were more surprising theoretically than empirically. And the reason, as Jenkins and Stanley note, was this:

A straight S-R interpretation, that reward strengthens a response and omission of reward weakens it, would predict a weaker response with partial reinforcement than with 100% reinforcement in acquisition, performance, and extinction. The facts of post-partial extinction are clearly opposed to this position [in one important respect] (10, p. 222).

Numerous attempts have been made both to modify reinforcement theory so as to make it account for such findings and, by its opponents, to discredit the theory. No systematic attempt will be made here to cite the resulting literature. Instead, the writer will sketch what seems to be the currently most justifiable interpretation of this phenomenon.

Following related suggestions by Skinner (19) and Humphreys (9), Mowrer and Jones (15) in 1945 advanced what they termed the "discrimination hypothesis" to account for the effects of partial reinforcement upon resistance to extinction. To quote:

If, during acquisition, a response (conceived as a more or less isolated movement) occurs repeatedly but is rewarded only now and then, the transition from acquisition to extinction will not be discriminated as sharply as if acquisition has involved reward for each and every response. With "faith" thus established that failure will ultimately be followed by success, "discouragement" is slower to set in (ergo, greater "resistance to extinction") when there is a change in objective conditions from acquisition (occasional reward) to extinction (no reward whatever) . . . (15, p. 309).

Or, as formulated elsewhere, the notion would be that ". . . resistance to

extinction is greater following intermittent reinforcement because the subject has difficulty in discovering that the extinction procedure *is* extinction, rather than just a continuation of positive training" (14).

One may thus say that the greater resistance to extinction which occurs following partial reinforcement is due to the fact that the subject is "fooled" or "tricked" into misperceiving as a continuation of acquisition (ultimate reward likely) what is in reality extinction (no possibility of reward). This type of thinking has received experimental confirmation in a study reported by Bitterman, Feddersen, and Tyler (1). However, it is cast in a terminology that is frankly "cognitive" and needs, if possible, to be translated into more objective language. This, fortunately, is not difficult to do.

Let us begin by considering the following situation. A hungry rat is put into an experimental set-up in which each depression of a bar causes a pellet of food to appear in a nearby trough. After the bar-pressing response is well stabilized, with 100% reinforcement, the procedure is altered so that bar pressing suddenly ceases to produce food until a minute has elapsed. As a result of making several "extinction," i.e., unrewarded, responses, two things happen to the rat. Each unrewarded response leads to some decrement in response strength, but also—and for present purposes, more importantly—each unsuccessful response *becomes an event which precedes the reinforcement which is going to come at the end of the minute*. For this state of affairs to obtain, it should be noted, the response strength must be great enough that the response will continue to occur from time to time throughout the minute interval—that is to say, the shift from 100% reinforcement to less-than-100% reinforcement must not be so drastic that the response will completely

extinguish before the possibility of the next reinforcement is at hand.

Another way of expressing this position is as follows. During 100% reinforcement a number of stimulus events have been the occasion upon which food is received. These have included the stimuli arising from the bar and its vicinity, the proprioceptive consequences of mounting and beginning to press the bar, and, especially, the stimuli arising from eating the food pellet which was obtained on the just-previous press. During intermittent training, however, an additional element is added to this complex of stimuli which set the occasion for reinforcement—namely, *the stimulus consequences of a number of previous bar-presses which have not been reinforced*. It is not surprising, then, that this discrimination should continue to operate for some time into the extinction session. And, particularly, if *many* unreinforced bar presses have been experienced during acquisition, it is quite reasonable to suppose that the discriminative control would hold up *far* into the extinction session before breaking down.

This argument has been developed in greater detail by Skinner (20), who has shown that the form of the curves which are generated under different kinds of schedules, both during acquisition and during extinction, can be readily accounted for by such assumptions as these.

Another important implication of the discrimination hypothesis, which has not often been taken into consideration by theorists in dealing with partial reinforcement, should be stressed. The implication is this: *Behavior will be maintained without decrement under a particular intermittent schedule only if the change from 100% reinforcement to less-than-100% reinforcement is gradual enough to maintain discriminative control*. For example, if an animal is

switched from continuous reinforcement to a fixed ratio of 1:100, the response will extinguish. However, going upward in small steps—say from 1:1 to 1:10 to 1:20, etc., and eventually to 1:100, will maintain the performance.

The implications of this well-substantiated empirical finding are rather far-reaching, especially when considered in relation to the usual kind of formulations of partial reinforcement and of probability learning. It would seem that frequency of reinforcement is not simply a parameter which has an orderly relation to "habit strength" whenever it is introduced into an experimental session. Rather, the effects of given frequencies of reinforcement will depend critically upon the manner in which the intermittency is introduced into the session. At least this seems to be true for the free-responding situation.

In other words, no generalizations can be made about given frequencies of reinforcement, per se, without further specification of the initial and boundary conditions of the case. A curve showing "habit strength" as a function of frequency of reinforcement would be meaningless, since a given frequency might or might not maintain performance, depending upon the conditions under which it were approached. Introduction of a frequency, that is to say, is not "independent of path" in the mathematical sense.

### III. INTERMITTENCY AND SECONDARY REINFORCEMENT

As originally expressed by Hull, the principle of secondary reinforcement was simple and straightforward: A stimulus which occurs "consistently and repeatedly" in conjunction with a reinforcing state of affairs acquires, itself, reinforcing potential. As will become clear in the next section, the simplicity of this principle is elusive; and certain



amendments will be necessary in order to account for all the relevant data. Once again the paradigm of *discrimination* will prove useful.

Before pursuing this line of thinking, however, let us return to the secondary reinforcement procedure reported in the first section.

In secondary reinforcement, we have seen, there are two kinds of "conditioning" or "association" processes going on, namely: (a) association of the potential secondary reinforcer with primary reinforcement; and (b) subsequent presentation of this secondary reinforcer as a consequence of some selected response. Theorists have not always clearly distinguished between these two processes. Although drawing the distinction, Hull (8) stressed only the second and neglected the first. Whereas he offered much mathematical machinery in order to deal with the strength of S-R connections as a function of number of reinforcements and other variables, Hull said relatively little about the manner in which the strength of secondary reinforcers builds up and decays. Although a "repeated and consistent" association was held to be necessary, the form of the relationship here was not indicated with the same explicitness as in the case of other variables.<sup>3</sup>

<sup>3</sup> Another ambiguity, which may turn out to be of the highest importance, concerns the way in which the "strength" of a secondary reinforcer is to be measured. In one sense this "strength" might be presumed to vary along the same dimensions as does amount and quality of a reinforcing agent, that is, by leading to a *stronger habit* when presented a given number of times as a consequence of responding on the part of the animal. Another indication of the potency of the secondary reinforcer is the extent to which behavior can be established and maintained by its presentation in the absence of primary reinforcement, i.e., how long it lasts before wearing out. Subsequent analysis may, in fact, indicate that variation on this latter dimension may occur within wide limits, whereas there may be no variation

In the preceding section we have attempted to explain why intermittent reinforcement of responses results in greater resistance to extinction. There remains, however, the problem of dealing with intermittency in the association of S<sup>r</sup> with primary reinforcement (or in Hull's terms, a repeated but *inconsistent* association), which is something quite different.

An example from human behavior may help to clarify this point. When the telephone rings, we are not always reinforced for answering it. But when it rings, we predictably answer. This kind of intermittency of reinforcement is to be distinguished from that in which a given response occurs "spontaneously," over and over again, and is only occasionally reinforced. Whereas much is known experimentally about the effect of intermittency in the case of nondiscriminative, freely occurring behavior, the effects of intermittency in the case of discriminative behavior is only just beginning to be studied. As the research findings reported in this paper indicate, discriminative behavior, with intermittent reinforcement, can be maintained at high strength. But the upper limit of the ratio of reinforced to nonreinforced occasions which is here possible is probably not so high as in the case where the responses occur freely (ratios as high as 1/100 or 1/1000, and even higher, have been reported).

The intermittent secondary-reinforcement technique, as described above, is analogous. We are dealing with intermittent reinforcement of a response (running to the water aperture), but by the addition of a discriminative stimulus—the buzzer—we now have a situation in which the *experimenter can reliably and predictably control the response by virtue of his ability to manipu-*

at all on the first dimension. Discussion in later sections should clarify this point.



*late the buzzer presentations.* The event then may be made a consequence of prior behavior on the part of the animal (bar pressing), and the experimenter can observe the effect on the frequency of this behavior. By using the variable ratio procedure, stability of behavior following the buzzer is insured. Long runs of  $S^r$  occurrences with no water immediately forthcoming will have been encountered during training; and when extinction (a shift from occasional reinforcement to *no* reinforcement) is begun, the change will not be very great. This point will be discussed in greater detail below, but we may anticipate it by saying that this specific response—approaching the water aperture—may itself be a critical event in determining the reinforcing properties of  $S^r$ . Any procedure which insures that this response “lasts longer” (remains under the discriminative control of the signal) during the later test procedure will then augment the secondary reinforcing potential of the signal.

Granted that the secondary reinforcer will last for some time during the test session—that is to say, that it is “good for” a certain number of presentations following the desired new behavior—there is still the problem of maximizing the work output of the animal. It is here that the *second* intermittent procedure—the familiar intermittent reinforcement of a *response* (to which the discrimination hypothesis was originally designed to apply)—is employed. By presenting  $S^r$  intermittently during the test session, the magnitude of the reinforcing effect is increased in three ways. (a)  $S^r$  presentations are stretched out over a longer period of time, and more work is required of the animal for each of the available  $S^r$  presentations. (b)  $S^r$  presentations tend to be distributed, rather than massed, and for this reason also the reinforcing effect may diminish more slowly. (c) The resistance to ex-

tingtion of the new response is increased, so that when  $S^r$  is itself withdrawn, the animal may nevertheless continue to perform the new response for a considerable period of time thereafter.

We are thus dealing with a controlling event,  $S^r$ , which is diminishing in potency at the same time that it is being utilized. Each presentation of  $S^r$  strengthens some behavior, and also leads to a decrement in its effectiveness on the next presentation. The problem is to use  $S^r$  judiciously while it is effective. We might conceive of an experimenter with a supply of hypothetical food pellets which, by some magical means, grow smaller and smaller with each presentation to the animal, until they finally reach microscopic size, or disappear entirely. If the experimenter faces the problem of getting as much work out of the animal as possible, over as long a period of time as possible, an intermittent reinforcement procedure would clearly be in order.

#### IV. SECONDARY REINFORCEMENT AND DISCRIMINATION FURTHER CONSIDERED (THE DINSMOOR HYPOTHESIS)

Following suggestions previously put forward by Schoenfeld, Antonitis, and Bersh (17), Dinsmoor (3) in 1950 advanced the thesis that secondary reinforcement and discrimination, in some kinds of situations, are identical—or, at least, that any stimulus which is a secondary reinforcer also has cue properties, and vice versa. The great importance of this thesis—and Dinsmoor reported experimental results which substantially confirmed it—is only just beginning to be appreciated. One of the major implications is this: A stimulus which is “repeatedly and consistently” paired with primary reinforcement presumably will *not* acquire secondary reinforcing properties unless, in addition, it has first played a discriminative role

of some kind, as a result of a differential reinforcement procedure.

On the basis of his experiment, Dinsmoor concludes:

It no longer seems to be necessary to provide separate theoretical accounts of the manner in which a discriminative or reinforcing stimulus gains or loses the power to raise the rate of a response. The previous distinction between the two types of stimulus appears to be reduced to a distinction between two categories of temporal schedule for the administration of the stimulus. This distinction is relevant chiefly to the determination of the distribution of the stimulus effects among the variety of responses in the organism's repertoire (3, p. 471).

What Dinsmoor means by the latter remark, apparently, is this: If a stimulus appears before a response which receives primary reinforcement (sets the occasion for a response), we call that stimulus a *cue* ( $S^D$ ); whereas, if the same stimulus is produced, not by the experimenter, but by some distinctive action on the part of the subject, we refer to it as a *secondary reinforcer* ( $S^r$ ) because it increases the likelihood of recurrence of this response. Said otherwise, a secondary reinforcer is a cue which the subject himself provides, by means of his own behavior. And when a stimulus gains or loses one capacity, it at the same time gains or loses the other.

The question now naturally arises as to how these considerations bear upon the so-called "discrimination" hypothesis previously discussed, in Section II. The relationship seems to be straightforward and consistent. According to the Dinsmoor thesis, a secondary reinforcer is a stimulus which enables the subject to discriminate between the conditions ( $S^A$ ) under which a particular response, if it occurs, will *not* be reinforced and conditions ( $S^D$ ) under which this response *will* be reinforced. And the secondary reinforcement capacity of such a stimulus is shown by the fact

that the subject will learn to make a *new* response if that response produces the cue stimulus,  $S^D$ .

Now, in a partial-reinforcement training procedure, a special condition is introduced. Here  $S^A$  (absence of  $S^D$ ) still means that the response in question will not "work"; i.e., in the situation described in Section I, the rat never gets water as a result of going to the water aperture when the buzzer has not sounded. But there is this difference: although the buzzer is thus a cue—which must be responded to (lest water be available, but not discovered, as it were)—it is an inconsistent one. As has been mentioned above, this inconsistency of reinforcement does not diminish the discriminative control which is exerted by the stimulus. And because it has been thus intermittently reinforced during acquisition, the lack of reinforcement is, so to say, no novelty when it occurs during extinction.

#### V. IS A SECONDARY REINFORCER MORE EFFECTIVE IF IT "RELEASES" BEHAVIOR?

The Dinsmoor hypothesis points to an equivalence between two functions of a stimulus which were formerly held to be separate. Although the results here are not as yet unequivocal, enough evidence now exists to justify the use of a discrimination training procedure whenever an effort is being made to give a stimulus secondary reinforcing properties. One may note these results, however, and still ask the question, Why? We must now examine briefly the possibility that the equivalence between discriminative stimuli and secondary reinforcing stimuli may be derived from more basic principles. Three possibilities suggest themselves.

1. In watching the behavior of the animals during the procedure described in Section I, one is reminded of Guthrie's



lastness-of-response hypothesis (7), or the postremity principle, as Voeks (24) has called it. Certainly a striking thing about the buzzer is the promptness and reliability with which it calls the subject away from whatever else it is doing and causes it to go to the water aperture. And when the bar is made available and so connected that the subject, by depressing it, can produce the buzzer (at least intermittently), one might argue that the response is preserved, or "protected." That is, the buzzer guarantees that the *last* response made will be one of bar pressing, since the buzzer then comes on and "takes the subject out of the situation."

A similar interpretation of the effects of reinforcers (both primary and secondary) has been suggested by Verplanck (23), and Estes (6). Following this kind of thinking in connection with *secondary* reinforcement, it is easy to see why the Dinsmoor restriction—that  $S^r$  must also be  $S^D$ —would be essential. When, following training, the stimulus is presented as a consequence of a response, it will elicit the behavior which it has formerly controlled in its discriminative role—in the present context, approaching the magazine. This uniform behavior, as long as it lasts, will "take the animal out of the situation," just as does a primary reinforcer. The Dinsmoor hypothesis may, in the light of this interpretation, be restated in somewhat stronger form: *A stimulus will be effective as a secondary reinforcer for new behavior if and only if it has some response already conditioned to it.* The secondary reinforcer, that is to say, must predictably "release" some behavior as a result of a prior training procedure; and when it then is used to strengthen new behavior, it is this response which is "released," not the stimulus *qua* stimulus which is the critical event. The question "How can secondary reinforcers be made to last longer?" may then be rephrased,

"How can discriminative stimuli be made to retain control over uniform behavior longer?"

Whether such specific control is both *necessary* and *sufficient* for reinforcement is a question which only further research can answer. This position would seem to be demanded by strict contiguity theorists such as Guthrie.

2. The former, weaker hypothesis—that discrimination is necessary, but not sufficient—could be readily assimilated by a variety of current theories of learning. A position such as that of Spence (22) or Seward (18), for example, could hold that the effect of a stable magazine-approach response depends upon  $r_g$ 's which are consistently elicited and protected from extinguishing.

3. Yet another possibility suggests itself. Because the buzzer reliably elicits running to the water aperture, the proprioceptive and other feedback from this response also become effective and the *range* of stimulation received is thereby enlarged. Hence, when the subject is later given an opportunity to press the bar, it is plausible to hold that there is now, so to say, a *double* course of secondary reinforcement: the buzzer (immediate feedback from bar pressing) and the stimulation produced by the behavior which the buzzer releases (mediated feedback from bar pressing). Thus, to generalize, one might conclude that any procedure where  $S^r$  is also an  $S^D$  will result in more potent secondary reinforcement because here there are more varied and more numerous stimuli provided by the mediating behavior. Perhaps much the same effect could be achieved by having  $S^r$  consist, not of a single, segmental stimulus change, but of several such stimuli. The latter might be more properly termed a *change in the situation*, thus being more nearly comparable to the change brought about by the animal's own response to  $S^r$  and the resulting stimulus consequences. This



is, however, at best conjectural; and the whole question of precisely why a response to an  $S^r$  makes it so much more efficient as a secondary reinforcer is still far from settled.

## VI. SUMMARY AND CONCLUDING REMARKS

Mowrer (13) has recently presented a theoretical interpretation of learning which depends critically upon the principle of secondary reinforcement. This theory is, in essence, a searching analysis and reformulation of the concept of *habit*, which is so ubiquitous in contemporary learning theory. Secondary reinforcement, as an experimental paradigm, is used as a major explanatory principle.

In the past, one of the disconcerting things about secondary reinforcement has been the extremely transitory and unstable nature of the demonstrated laboratory effect. One thing seems clear: a theory which leans so heavily upon this principle is doomed at the outset if the raw empirical phenomenon cannot be demonstrated in more potent form. Before using the paradigm as an explanatory device to account for (reduce to more basic principles) the most complex forms of mammalian behavior, there seems to be the anterior question as to whether, in fact, secondary reinforcement is *real*, qua phenomenon. Although experimental work within the last three decades has provided suggestive evidence, the results have been far from unequivocal. Actual experimental demonstration—such that a disinterested observer could view the animal in question and be convinced that “something is really happening”—has been lacking.

It has been the purpose of this paper to discuss a method by which secondary reinforcement can be made highly effective and stable, thus giving the needed empirical buttressing to the kind of

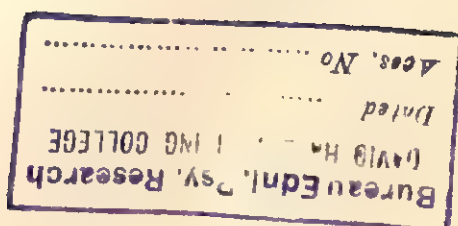
theorizing cited. This method depends upon the use of *intermittent* reinforcement, thereby also further accenting the very great importance of the intermittent reinforcement technique for psychological theory.

## REFERENCES

1. BITTERMAN, M. E., FEDDERSON, W. E., & TYLER, D. W. Secondary reinforcement and the discrimination hypothesis. *Amer. J. Psychol.*, 1953, 66, 456-464.
2. CLAYTON, F. L. Secondary reinforcement as a function of reinforcement scheduling. *Psychol. Reports*, 1956, 2, 377-380.
3. DINSMOOR, J. A. A quantitative comparison of the discriminative and reinforcing functions of a stimulus. *J. exp. Psychol.*, 1950, 40, 458-472.
4. DINSMOOR, J. A. The effect of periodic reinforcement of bar pressing in the presence of a discriminative stimulus. *J. comp. physiol. Psychol.*, 1951, 44, 354-361.
5. DINSMOOR, J. A. Resistance to extinction following periodic reinforcement in the presence of a discriminative stimulus. *J. comp. physiol. Psychol.*, 1952, 45, 31-35.
6. ESTES, W. K. Generalization of secondary reinforcement from the primary drive. *J. comp. physiol. Psychol.*, 1949, 42, 286-295.
7. GUTHRIE, E. R. *The psychology of learning*. (Rev. ed.) New York: Harper, 1952.
8. HULL, C. L. *Principles of behavior*. New York: Appleton-Century-Crofts, 1943.
9. HUMPHREYS, L. J. The effect of random alternation of reinforcement on the acquisition and extinction of conditioned eyelid reactions. *J. exp. Psychol.*, 1939, 25, 141-158.
10. JENKINS, W. O., & STANLEY, J. C., JR. Partial reinforcement: a review and critique. *Psychol. Bull.*, 1950, 47, 193-234.
11. McCLELLAND, D. C., ATKINSON, J. W., CLARK, R. A., & LOWELL, E. L. *The achievement motive*. New York: Appleton-Century-Crofts, 1953.
12. MELCHING, W. H. The acquired reward value of an intermittently presented neu-

- tral stimulus. *J. comp. physiol. Psychol.*, 1954, 47, 370-374.
13. MOWRER, O. H. Two-factor learning theory reconsidered, with special reference to secondary reinforcement and the concept of habit. *Psychol. Rev.*, 1956, 63, 114-128.
  14. MOWRER, O. H. Learning theory and behavior. Urbana, Ill.: Author, 1957.
  15. MOWRER, O. H., & JONES, H. M. Habit strength as a function of pattern of reinforcement. *J. exp. Psychol.*, 1945, 35, 293-311.
  16. SALTZMAN, I. J. Maze learning in the absence of primary reinforcement: a study of secondary reinforcement. *J. comp. physiol. Psychol.*, 1949, 42, 161-173.
  17. SCHOENFELD, W. N., ANTONITIS, J. J., & BERSH, P. J. A preliminary study of training conditions necessary for secondary reinforcement. *J. exp. Psychol.*, 1950, 40, 40-45.
  18. SEWARD, J. P. How are motives learned? *Psychol. Rev.*, 1953, 60, 99-110.
  19. SKINNER, B. F. *The behavior of organisms*. New York: Appleton-Century, 1938.
  20. SKINNER, B. F. Are theories of learning necessary? *Psychol. Rev.*, 1950, 57, 193-216.
  21. SPENCE, K. W. The role of secondary reinforcement in delayed reward learning. *Psychol. Rev.*, 1947, 54, 1-8.
  22. SPENCE, K. W. *Behavior theory and conditioning*. New Haven: Yale Univer. Press, 1956.
  23. VERPLANCK, W. S. The operant conditioning of human motor behavior. *Psychol. Bull.*, 1956, 53, 70-83.
  24. VOEKS, V. W. Formalization and clarification of a theory of learning. *J. Psychol.*, 1950, 30, 341-362.

(Received February 25, 1957)



## AN OPPONENT-PROCESS THEORY OF COLOR VISION

LEO M. HURVICH<sup>1</sup> AND DOROTHEA JAMESON<sup>1</sup>

*Eastman Kodak Company*

The two major theoretical accounts of color vision are those classified as the Young-Helmholtz and the Hering types of theories. For many years the former has been judged by most workers in the field to provide the simplest explanation of the way in which light stimuli give rise to color sensations. The advantages that appear to favor the Young-Helmholtz three-component hypothesis are two: it is parsimonious, and its postulates are easily quantifiable and hence subject to precise experimental test. In its parsimonious and easily quantifiable form, the theory is simple: in addition to the rods which subserve twilight vision, the eye contains three kinds of cone photoreceptors; each type of cone contains a differently selective photochemical substance; each is associated with its own specific nerve fiber; and each cone-photochemical-nerve fiber system is correlated with one of the three specific "fundamental" color sensations, namely, red, green, and blue (or violet). All sensations are considered as compounded of varying amounts of these three excitatory systems, with white arising from equal and simultaneous excitation of all three, and yellow from equal red and green excitations.

The Young-Helmholtz three-cone, three-nerve, three-sensation theory derives directly from the basic fact of color mixture, namely, that all visible hues can be matched by the mixture, in proper proportions, of only three physical light stimuli. Based squarely on this fact, the theory is readily quantified in terms of the three measurable variables of color mixture experiments. But the three measured variables, it

must be emphasized, are the three physical light stimuli used in the color mixture experiments; they are not the postulated three "fundamental" color sensations, for with each different stimulus triad used for color matching a different and equally valid triad of color mixture functions is obtained. Consequently, throughout some hundred years since the original formulation of the idea, a continued series of attempts has been made to find the proper transformation of the three measured color-mixture curves that will bridge the gap and yield the unique spectral distribution curves of the desired physiological correlates of the three postulated "fundamental" color sensations. An infinity of such transformations is available for trial, and almost every serious adherent of the theory has proposed at least one new set of "fundamental sensation curves" (48, pp. 368-372). The search, however, continues, because serious defects have been found in every proposal made thus far. When the explanatory or predictive power of the theory in any given quantified form is tested it cannot handle more than a limited number of facts satisfactorily (11, p. 805).

Moreover, some facts of color experience seem unassimilable into the framework of the simple Young-Helmholtz theory with its three independent, fundamental, process-sensation systems. How can this system of three independent processes be made to account, for example, for the apparent linkages that seem to occur between specific pairs of colors as either the stimulus conditions or the conditions of the human observer are varied? Why should the red and green hues in the spectrum predominate

<sup>1</sup> Now at New York University.



at low stimulus levels, and the yellow and blue hue components increase concomitantly as the spectrum is increased in luminance (43)? Why, as stimulus size is greatly decreased, should discrimination between yellow and blue hues become progressively worse than that between red and green (4, 10)? Why should the hues drop out in pairs in instances of congenital color defect, or when the visual system is impaired by disease (29, 31)? On the other hand, since the sensation of white is granted no special physiological process in this parsimonious theory, but occurs as the fusion product of three equally large fundamental hue sensations, how account for the large degree of independence of white and chromatic qualities when the adaptation of the visual system is varied (37, 41)?

As more and more *ad hoc* hypotheses are added to the original Young-Helmholtz formulation in order to answer these and other problems forced by the increasing accumulation of experimental data, we naturally find the formulation becoming less and less precise and quantifiable, and obviously less parsimonious. We also find, however, that exactly those phenomena that require modification and extension of the simple "three-color theory" remind us more and more of its chief theoretical rival, the Hering theory of three paired, opponent color processes.

In view of this situation, it seems highly desirable that we take a close second look at Hering's alternative approach to an understanding of color phenomena. The vast accumulation of psychophysical data for which any adequate theoretical proposal must account requires that the basic postulates of the theory, as outlined qualitatively by Hering (13, 14), be restated in quantitative terms for such a critical scrutiny to be most meaningful. This paper will review our attempt to provide such a

quantitative restatement, and will summarize briefly some of the critical comparisons between the theoretical deductions and relevant psychophysical data. (Detailed quantitative accounts are given in 21, 22, 23, 25, 26, 27.)

## BASIC SCHEMA FOR THE HERING THEORY

### *The Three Variables*

The Hering theory is like the Young-Helmholtz theory in that it, too, postulates three independent variables as the basis for color vision, but the Hering variables are three pairs of visual processes directly associated with three pairs of unique sensory qualities. The two members of each pair are opponent, both in terms of the opposite nature of the assumed physiological processes and in terms of the mutually exclusive sensory qualities. These paired and opponent visual qualities are yellow-blue, red-green, and white-black.

The basic schema for the opponent-colors mechanism is shown diagrammatically in Fig. 1. The three paired opponent response systems are labeled *y-b*, *r-g*, and *w-bk*. The convention of positive and negative signs is used to

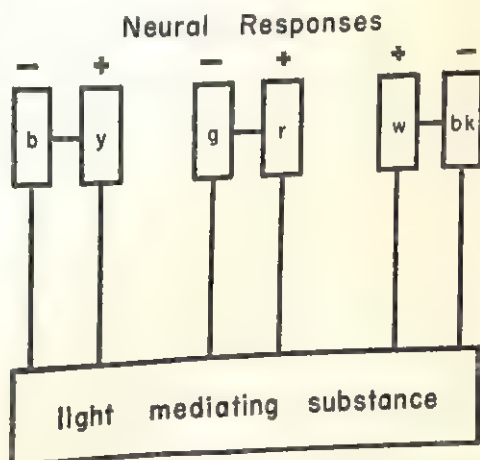


FIG. 1. Basic schema for Hering opponent-colors mechanism.

indicate that each neural system is capable of two modes of response that are physiologically opponent in nature, and that the paired sensory qualities correlated with these opposed modes of response are also mutually opponent or exclusive. That is, we may experience red-blues or green-blues but never yellow-blues, and we see yellow-greens or blue-greens, but never red-greens, and so on. In the absence of any external visual stimulus, the state of the visual system is assumed to be a condition of active equilibrium, and this equilibrium condition is associated with the neutral, homeogeneous "gray" sensation perceived after a long stay in complete darkness. This sensation is quite different from the black experience of the white-black opponent pair. Blackness arises neither by direct light stimulation nor in the simple absence of light, but rather by way of either simultaneous or successive contrast during, or following, light stimulation of some part of the retina.

### *Properties of Paired Systems*

The three pairs of visual response processes are independent of each other; that is, they have different response thresholds, they follow different laws of increase with increase in strength of stimulation, and probably have different time constants. The achromatic system is the most sensitive; that is, the amount of photochemical absorption necessary to excite the achromatic white response is less than the amount of photochemical activity required to stimulate either the  $y-b$  or  $r-g$  chromatic pairs. This characteristic accounts for the existence of the so-called achromatic interval, i.e., the fact that spectral lines appear achromatic at the absolute threshold for visibility (42, p. 167). Similarly, the red-green system has a lower threshold than the yellow-blue one. The failure of the yellow-blue system to respond at near-

threshold levels that are sufficient to activate the red-green system exhibits itself in the facts of so-called "small field dichromasy," in which the eye behaves, with respect to stimuli that are very small in area as well as of low intensity, in a manner similar to the congenital tritanope, i.e., a specific type of "color blind" individual for whom yellow and blue discriminations are impossible and the only hues seen are reds and greens (4, 49).

With increase in level of stimulation the different paired systems also show differences in rate of response increase, such that the achromatic response increase is probably the most rapid of the three, with the result that at very high intensities all spectral stimuli show a strong whitening, or desaturation, relative to their appearance at some intermediate luminance level (42, p. 168). Of the two chromatic pairs, the yellow-blue system, although exhibiting a higher threshold, shows a more rapid rate of increase in response with increase in luminance than does the red-green system. Thus, the mixed hues of the spectrum—the violets, blue-greens, yellow-greens, and the oranges—all vary systematically with increase in spectral luminance, and all show a tendency to be more blue or yellow, respectively, at high luminances, and more red or green at the lower luminance levels (the Bezold-Brücke hue shift phenomenon).

The opponent systems show a tendency toward restoring the balanced equilibrium condition associated with the neutral "gray" sensation. Thus excitation, say, of the  $r$  process in the  $r-g$  system results in a decrease with time in  $r$  responsiveness, and in an increase in the responsiveness of the opponent  $g$  process. If we think of the  $r$  process as perhaps associated with the building up of an electrical potential in the neural system, and of the  $g$  process as associated with the collapse of the po-



tential during impulse firing, then it is easy to see that as the neural potential is increased to higher values there will be a tendency to resist further build up, and also an increased disposition of the tissue toward impulse firing in order to restore the potential to its normal equilibrium value. Although we are not at all ready to ascribe a specific neural correlate of this sort to the postulated opponent processes at this time, the neurophysiological parallels are useful for conceptualizing the opponent-process notion as a real biological phenomenon.

To return to our example, if the responsiveness of the opponent  $g$  process tends to increase as  $r$  excitation is continued, then when the  $r$  stimulus is removed we can expect  $g$  activity to be released, strongly at first, then more slowly, and ultimately fading out as equilibrium is again approached. The sensory correlate of this reversal of opponent activities with removal of stimulation is, of course, the familiar phenomenon of the complementary after-image. If the stimulus (of constant magnitude) is not removed but continues to act for a considerable length of time, then the  $r$  process, whose responsiveness is being continuously decreased, will eventually cease to respond further, and a new equilibrium state will be reached. The disappearance of a sensory response with continued constant stimulation can be observed either by the *Ganzfeld* technique, in which the whole retina is uniformly illuminated by diffuse light (18), or by the "painted image" technique, in which optical means are used to fix a well defined image on the retina in such a way that its retinal position remains constant and independent of eye movements (39). By either method the eventual result of continued steady stimulation is a disappearance of the visual experience: the light seems to have gone out in the

*Ganzfeld* situation, or, in the fixed-image situation, the perceived object simply fades out of view.

Not only are the visual responses modified by changes in time in the excitabilities of the opponent processes, but they are also importantly affected by spatial interaction among the various elements of the visual field. Within certain limits there is evidence of summation of similar kinds of activity in adjacent elements, as in threshold responses for small stimulus areas (5, pp. 846-852). But perhaps more important for the over-all functioning of the visual system are the antagonistic interactions, such that  $r$  activity in one area induces  $g$  activity in adjacent areas, and similarly for the yellow-blue and white-black paired response systems. These opponent spatial induction effects are evident in all the familiar color and brightness contrast phenomena (35, pp. 138-142). They are probably also primarily responsible for the great visual-image clarity that characterizes vision in spite of the fact that the optical system of the eye is obviously imperfect, and that consequently the light image formed on the retinal surface lacks sharply defined boundaries (17, pp. 151-159). The spatial interaction causing intensification of opponent qualities at adjacent surfaces would seem an ideal crispening device to sharpen up the initially blurred retinal image.

### *Photochemical Postulates*

In addition to the various temporal and spatial induction effects, which are assumed to be based in the neural visual-response tissue, visual adaptation probably also involves changes in the photochemical activities that initiate the neural responses, since a certain amount of photochemical bleaching is expected to occur with continued exposure of the photosensitive materials to a retinal light stimulus. In order



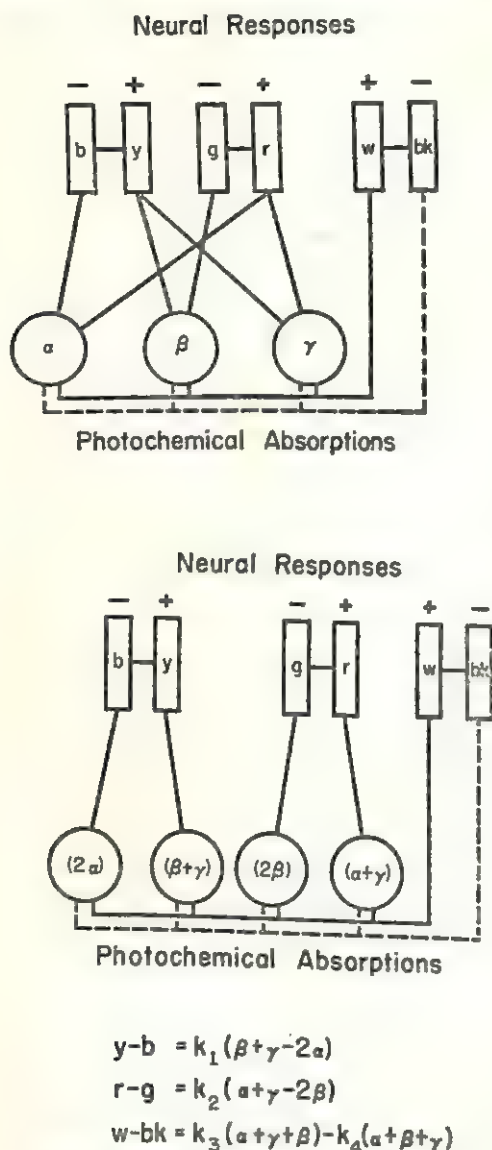


FIG. 2. Schematic diagram showing relations between photosensitive materials  $\alpha$ ,  $\beta$ , and  $\gamma$  and neural opponent response processes  $y-b$ ,  $r-g$ , and  $w-bk$ .

for the three paired opponent-response systems to be selectively stimulated, there must, of course, be more than one substance available for photochemical mediation between the incident light and the neural excitation. Whatever the specific nature of the photosensitive materials, they must form a link in the system of three independent variables,

and hence we have postulated three independent photosensitive materials, which we may call  $\alpha$ ,  $\beta$  and  $\gamma$ .

Our schematic model now takes the form shown in Fig. 2A or 2B. The three independent photosensitive materials may be contained in discrete retinal units with complex interconnections to the neural response systems, as shown in Fig. 2A, or two or more of these materials may be combined in receptor units having simpler connections to the neural response systems, as diagrammed in Fig. 2B. There is no way of differentiating these models in terms of visual behavior; and however the three photochemicals may be segregated or combined in the retina, and whatever the number of different photoreceptor units, there remain only three independent photosensitive materials, and the theory remains a three-variable, opponent-colors schema.

#### QUANTIFICATION OF OPPONENTS THEORY

Since our aim is to present this schema in quantitative terms, one of the first questions that has to be asked is this: Is it possible to obtain by psychophysical experiment direct measurements of the spectral distributions of the three basic response variables of the Hering theory?

#### Measures of Achromatic and Chromatic Responses

It can fairly be assumed that the achromatic, white response is closely connected with the distribution of the brightness quality throughout the visible spectrum, and Fig. 3 therefore shows two functions (which we have measured by a threshold technique) that give the whiteness distribution of an equal energy spectrum for two observers (20). The induced rather than directly stimulated black component of the achro-

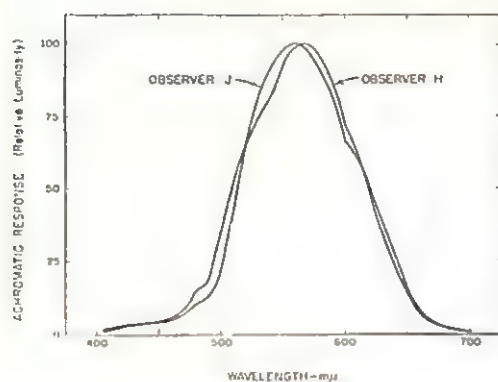


FIG. 3. Whiteness distribution of an equal energy spectrum for two observers.

matic white-black response pair has this same distribution, but of opposite sign, since the strength of the black contrast response is directly related to the magnitude of either the surrounding or the preceding whiteness or brightness.

A method for determining the spectral distributions of the paired chromatic responses is implicit in the opponents theory itself. Since the two members of each hue pair are mutually opponent or exclusive, then a yellow response of given strength should be exactly canceled by a stimulus that, taken alone, elicits the same magnitude of blue response, and a similar relation should hold between red and green responses. Thus a null method, based on the antagonism of the two members of each hue pair, can be used to measure the spectral distributions of the chromatic responses. In brief, a wave length is first selected that evokes, say, a blue hue response. The observer then views, in turn, a series of spectral wave lengths that appear yellowish in hue (yellow-greens, yellow, and yellow-reds). To each of these yellow stimuli just enough of the previously selected blue stimulus is then added exactly to cancel the yellow hue without introducing any blueness. The observer simply reports when the test field appears neither yellow nor blue; the hue remainder that he sees

may be green, neutral, or red, depending on the test wave length. Knowing the energies of the series of spectral yellow stimuli, and having determined experimentally the energy of the blue stimulus of fixed wave length that is required for the hue cancellation in each case, we can now plot the distribution of the relative magnitudes of yellow hue response evoked by the various test wave lengths. The procedure is simply reversed to obtain the distribution of

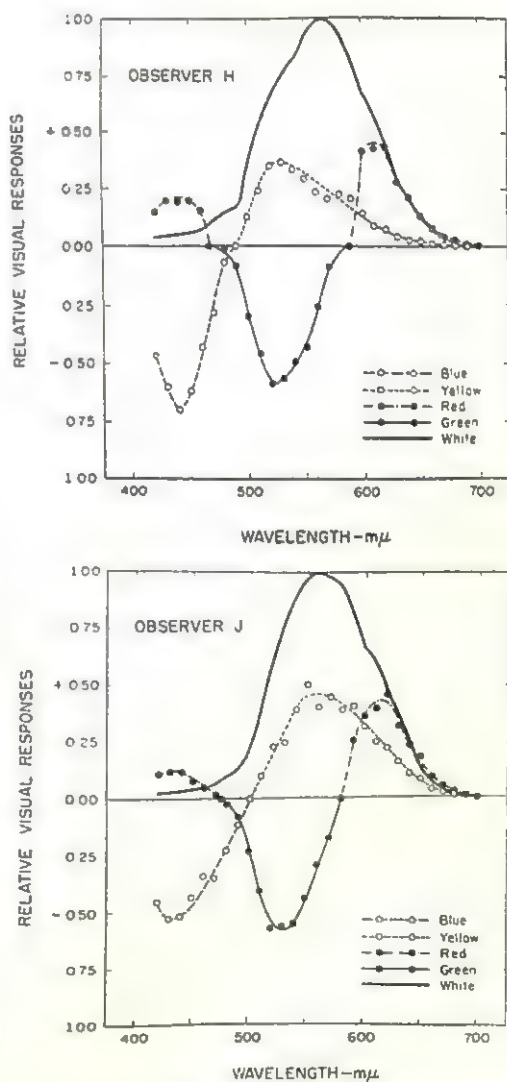


FIG. 4A and 4B. Chromatic and achromatic response functions for equal energy spectrum for two observers.

the blue component of the yellow-blue pair; that is, varying amounts of a fixed wave length of yellow hue are used to cancel the blue hue quality of a series of "blue" test wave lengths. By using a red stimulus of fixed wave length and variable energy to cancel the greens, and a green stimulus to cancel the reds, the spectral distribution of the red-green pair of chromatic responses is similarly determined.

Two sets of paired chromatic response vs. wave length functions that were measured in this way (25), together with the achromatic spectral functions shown in Fig. 3, are plotted in Fig. 4 for an equal energy spectrum. The opponent members of each hue pair have been given arbitrary positive and negative designations, to correspond with their opponent characteristics. Thus the positive values of the red-green function indicate redness, and the negative values greenness. Similarly, the positive values of the yellow-blue function indicate yellowness, and the negative values blueness.

These are the psychophysical functions that represent the spectral distributions of the three independent variables of the Hering opponent-colors theory for these two observers. They are assumed to be directly correlated with the response activity of the visual nervous tissue (retina, optic nerve, and visual centers), and should not be taken as photochemical absorption spectra, about which these data tell us nothing.

### *Brightness, Hue, and Saturation*

The psychophysical opponent-response functions shown in Fig. 4 provide a direct description of the appearance of the spectrum, for these observers, for a neutral condition of bright adaptation and at the moderate level of luminance for which the functions were obtained. Thus, all wave lengths evoke some whiteness as well as hue; the whiteness

and brightness of an equal energy spectrum is relatively small at the two spectral extremes and relatively high at the intermediate wave lengths.

The short wave lengths appear as red-blue hues (violets); there is a narrow band of pure or unique blue where the red-green function is equal to zero; then come the blue-greens, followed by a narrow band of unique green at the wave length where the yellow-blue function is equal to zero; this is followed by the yellow-greens, and then pure yellow occurs at the second intersection of the red-green function with the zero ordinate value; and finally the yellow-red hues appear in the long wave length region (19). A quantitative expression for hue, a "hue coefficient," can be obtained by taking the value of one of the chromatic responses, say, the yellow value at 550  $m\mu$ , relative to the total of all chromatic responses at that wave length, in this case, yellow plus green.

The saturation of the color depends on the relative amounts of chromatic and achromatic responses. At the two spectral extremes where the chromatic responses are large relative to the white response, the spectral saturation is high. Where the reverse is true, spectral saturation is low. This can be expressed quantitatively in the form of a "saturation coefficient." To use the same example, the total of the yellow-plus-green values relative to the white plus yellow plus green is relatively low at 550  $m\mu$ , and this wave length appears much less saturated than does, say, either 440  $m\mu$  or 670  $m\mu$ .

### *Color Mixture*

Since color-mixture experiments simply involve matching the three perceived qualities evoked by one stimulus by the proper mixture of three other stimuli, it is possible to determine the color-mixture relations that are inherent in the response curves of Fig. 4 for



any three arbitrarily selected mixture primaries. That is, the red-green value, the yellow-blue value and the white value of the total visual response to any wave length of unit energy are matched by the totals of the three corresponding values for the three mixture primaries when the latter stimuli are combined in the proper ratios. On paper, the color equations for most spectral matches require the admission of negative values for one of the mixture primaries. In actual color-mixture experiments, these negative values are realized by removing one of the mixture primaries from the matching field and adding it to the test stimulus.

To calculate, for example, the amounts of energy required for a color match to a given wave length  $\lambda$  by the mixture of the spectral primaries 460 m $\mu$ , 530 m $\mu$  and 650 m $\mu$ , let  $a$  = the energy at 460 m $\mu$ ,  $b$  = the energy at 530 m $\mu$ , and  $c$  = the energy at 650 m $\mu$ . The three equations to be solved for these three unknowns  $a$ ,  $b$ , and  $c$  are then:

$$a(r_{460}) + b(r_{530}) + c(r_{650}) = r_{\lambda}$$

$$a(y_{460}) + b(y_{530}) + c(y_{650}) = y_{\lambda}$$

$$a(w_{460}) + b(w_{530}) + c(w_{650}) = w_{\lambda}$$

The values for  $r$  (or for  $-r$  when the response function is negative, indicating that the hue is green rather than red), for  $y$  (or for  $-y$  when the response is blue rather than yellow), and for  $w$  are then read from the response functions for unit energy for each wave length in question. (See Fig. 4.) The values  $r_{\lambda}$ ,  $y_{\lambda}$  and  $w_{\lambda}$  represent the unit energy response values for any spectral wave length for which a color-mixture equation is to be calculated. Solving this set of three equations for the three unknowns  $a$ ,  $b$ , and  $c$ , we then have a color-mixture equation of the form

$$a_{460} + b_{530} + c_{650} = 1_{\lambda}$$

This equation, which is expressed in energy units, may be converted to photo-

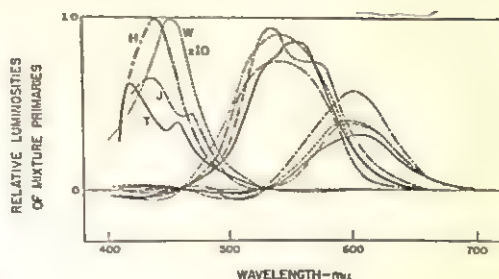


FIG. 5. Calculated color mixture functions for observers H and J and experimental color mixture functions for observers T and W (24, 47).

metric units in the usual way by multiplying each energy value by the relative luminosity (given by the achromatic response function) at the given wave length.

Color-mixture relations calculated in this manner for wave lengths  $\lambda$  from 420 m $\mu$  through 700 m $\mu$  from smoothed visual response data for two observers are shown in Fig. 5. The two additional sets of color-mixture functions (for the same three mixture primaries) that are shown for comparison in the figure are the results of actual color-mixture experiments by W. D. Wright and L. C. Thomson (24, 47).

Since the relations between the measured response functions and the color-mixture data are, as we have just seen, known for two individual observers, it is now also possible (by assuming specific spectral loci for the unique hues) to reverse the procedure and derive opponent-response functions from the color-mixture data for Wright and Thomson, or for any other observer whose color-mixture data are available. Since it seems preferable to develop a general theoretical model on the basis of a representative average, rather than a small number of individual observers, we have used for the model chromatic and achromatic response functions derived from the average color-mixture data for the CIE international standard observer (30). These derived functions

are shown in Fig. 6. (The details of the derivation are given in 22.) They are, of course, smoother and more regular than the individual, measured functions, but in other respects they are quite similar (compare Fig. 4).

### Photochemical Distributions

The specific set of  $\alpha$ ,  $\beta$ , and  $\gamma$  photosensitive absorption functions that have been assumed for the theoretical model are shown in Fig. 7. These curves have not been measured, and they have the particular forms shown and the high degree of overlap exhibited because of the specific interrelations that we have postulated a priori between the photochemical events and the neural response mechanisms. Once the photopigments actually present in the human retina have been identified by the biochemists, the visual theorist will have no need to make such a priori postulates, and the specific interrelations required between the identified photosensitive materials and the neural processes underlying the color responses can easily be deduced.

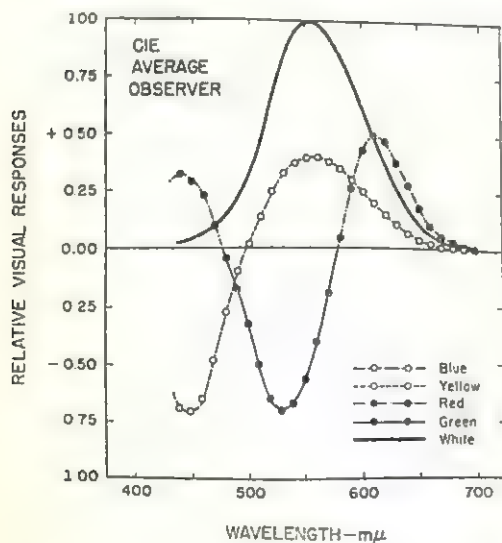


FIG. 6. Theoretical chromatic and achromatic response functions for equal energy spectrum for CIE average observer.

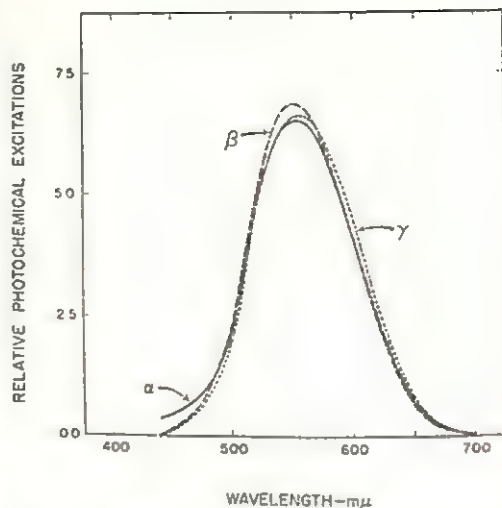


FIG. 7. Spectral distribution curves for assumed photosensitive materials.

As matters now stand, however, the functions shown in Fig. 7 meet the basic demands of the known facts, and any changes in these theoretical absorption functions that will no doubt be required by results of photochemical researches will not importantly affect any of the basic postulates of the theoretical model. The broadness and similarity of shape of all three selective functions that we have assumed are characteristic of all visual pigments so far identified in a variety of animal species (2).

These assumed photopigment distributions do not enter into the consideration of color phenomena, in normal vision, until we come to an examination of some of the phenomena of chromatic adaptation in which selective photochemical bleaching seems to act as one of the important determinants. The other determinants are, of course, the spatial and temporal induction effects in the neural opponent response processes that have been mentioned earlier.

### Dependence of Hue and Saturation on Both Wave Length and Luminance

What are the visual phenomena for which the model can account? As we

have already indicated, the measured chromatic and achromatic response functions provide a direct and quantifiable description of the color sensations evoked by any stimulus of specified wave-length composition (23). The achromatic, white function is taken as a direct expression of spectral brightness. Spectral hue, which is determined by the chromatic responses evoked by each wave length, can also be expressed quantitatively as a coefficient value relating the magnitude of response of one chromatic system to the total of all chromatic responses at that wave length. An example of such a hue coefficient function for a moderate level of luminance is shown in Fig. 8. It is clear, from the varying rate of change in the hue coefficient function from one region of the spectrum to the next, that an observer's ability to discriminate among neighboring wave lengths on the basis of hue changes alone will also differ for the different regions of the spectrum. This discriminative capacity is obviously also quantifiable in terms of the amount of wave-length change required to elicit a threshold change of fixed amount in the value of the hue coefficient. With change in the luminance at which the spectrum is presented, these coefficient

functions will be altered, in the sense that the yellow-blue values will increase at the higher luminances, and will be diminished at the lower luminances. This is so because, in accordance with the different energy-vs.-response function postulated for the yellow-blue system as compared with the red-green one, as the excitation level is increased, the yellow and blue spectral responses will be uniformly magnified relative to the red and green ones at the higher levels, and uniformly diminished at the lower levels. Although the exact differential between the two paired systems is not known, under certain circumstances an over-all difference in response magnitudes of approximately 20 per cent seems to occur for a log unit change in luminance. Thus, at some wave length for which, say, the red and yellow responses are equal at a luminance of 10 mL, the yellow will be about 20% greater than the red at 100 mL, and about 20% less at a luminance of only 1 mL. If we assume this 20% differential between *y-b* and *r-g* response magnitudes per log unit of luminance change as a reasonable value, and compute the spectral hue coefficients for a range of approximately three log units of luminance variation, then we can specify the amount of hue shift associated with a change in intensity of any wave length. Conversely, we can also specify the wave length changes necessary to maintain a constant hue sensation (constant hue coefficient value) as the luminance is increased or decreased. The latter procedure has been used to obtain the functions shown in Fig. 9, and the curves in the upper part of the figure are functions measured by Purdy in an actual experiment of this sort (38).

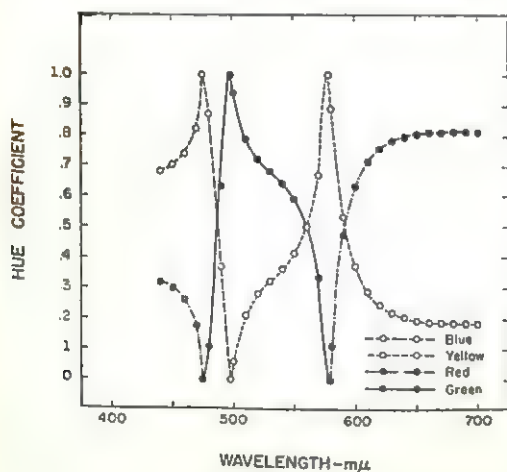


FIG. 8. Spectral hue coefficients. Moderate luminance.

These hue phenomena do not involve the achromatic response pair at all, and depend only on the two paired chromatic response systems. Whatever the



chromatic response to a given stimulus, the perceived color saturation clearly will also depend on the extent to which the white system is simultaneously responding. For any given amount of chromatic response, the color will obviously appear less saturated if there is a large magnitude of white response to dilute the color, and more saturated if the white component of the total response is relatively small. The perceived saturation of the spectrum is also expressed as a quantitative coefficient function. (See Fig. 10.) Here the value taken as the saturation coefficient is the ratio of the total chromatic to the chromatic-plus-white responses at each wave length. The relatively high values at the spectral extremes and the minimal value in the pure yellow region are perfectly consistent both with qualitative reports and with the experimental data on this problem (e.g., 28). Again, as in the hue functions, the rate of change of the saturation coefficient from one spectral region to the next is indicative of a varying discriminative capacity with respect to wave length; and, again, the form of the function as shown applies to a moderate luminance level and varies in a determinable man-

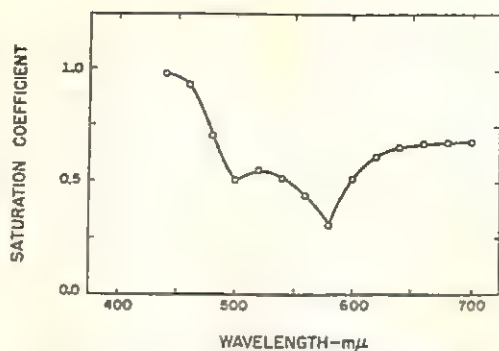


FIG. 10. Spectral saturation coefficients. Moderate luminance.

ner with change in the level of excitation in accordance with the different energy-vs.-response rates of the three independent response systems.

In view of the variations in the hue and saturation functions with change in luminance, we should expect that discrimination functions that depend on changes in these two color attributes, such as discrimination of one wave length from the next in an equal brightness spectrum, would also reflect such a dependence on luminance. Figure 11 shows, in the upper half, a set of wave-length discrimination functions obtained at two luminance levels by Weale (45).

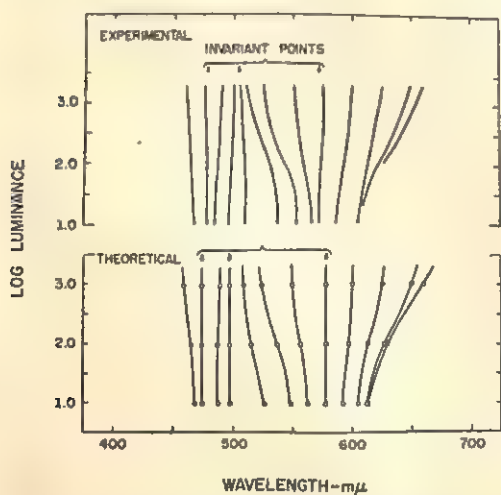


FIG. 9. Constant hue contours as measured by Purdy (38) and as predicted by theory.

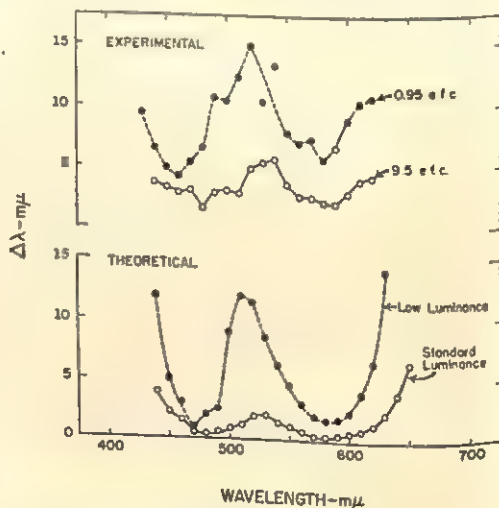


FIG. 11. Wave-length discrimination functions for two luminance levels as measured by Weale (45) and as predicted by theory.

The higher values of difference threshold obtained at the low luminance level may be explained by a general reduction of discriminative capacity in dim light. The shift of the midspectral maximum toward lower wave lengths, and the relatively greater heightening of the minimum in the yellow region, cannot, however, be attributed to such a generalized reduction in discriminatory capacity. The selectively greater loss in yellow and blue responses at the low-luminance level that is one of the postulates of our model does, however, account for changes of exactly this sort in the form of the function. This is shown by the two theoretical functions computed from pairs of spectral hue and saturation functions that are associated with the two specified luminance levels. Since brightness is kept constant in such experiments, only the hue and saturation variables need be considered in our analysis of these functions (22).

### *Chromatic Adaptation*

The phenomena that we have treated thus far all refer to the individual with normal color vision in a neutral state of adaptation. What of his color perception after the visual system has been exposed for some time to a strongly colored illuminant? For analytical purposes, the simplest situation of this sort is the one in which the eye has been exposed to a large surround field of given color and luminance, and the test stimuli are viewed at the same level of luminance as the surround. Under these circumstances, the three photochemical receptor substances will probably have undergone some selective bleaching, and because of the similar brightness of the surround and test fields, spatial induction effects in the neural response processes will probably be fairly constant. To simplify the treatment for these particular conditions, therefore, we may ignore the constant neural inductions

and consider the photosensitive changes as exercising a controlling influence on the response systems.

We know that under these circumstances the color-mixture data do not change. That is, with uniform chromatic adaptation, any change in the perceived color of one side of a bipartite color-mixture field will also occur on the other side, and to exactly the same extent. Thus a color equation that has been made with the eye adapted to a neutral white light will also be a valid equation when the eye is adapted to a colored illuminant (15). These important constancies of color equations mean that whatever photochemical changes occur with adaptation must occur in a very specific way. That is, the spectral distribution functions representing the three selective photochemicals may be selectively multiplied or reduced by a constant factor, but no one of them can change its form (44, pp. 211-212). In other words, any single substance cannot lose a greater percentage of its absorption at one wave length than it loses at another wave length. Thus, exposure to a colored light can cause any one of the postulated photochemical functions shown in Fig. 7 to be multiplied or divided by a constant amount, but this is the only alteration in the photosensitive functions that is consistent with the fact that color equations are invariant with chromatic adaptation.

The extent to which the three substances are selectively attenuated as a result of exposure to colored light is clearly controlled by the light stimulus itself. That substance which initially absorbs most of the adapting light will suffer the greatest relative bleaching, and the substance which absorbs relatively little of the adapting light will be relatively little affected by it. Thus, by determining their relative absorptions of the adapting light, we can compute the relative changes in the heights.

of the photosensitive distribution functions for the three photopigments that we have postulated. Since the excitations of the opponent response systems depend on these photochemical light absorptions (see Fig. 2), we can now also determine the forms and magnitudes of the chromatic and achromatic response functions for the new condition of adaptation. In spite of the close overlap of the photosensitive functions that we have postulated, the "adapted" chromatic response functions determined in this way change in striking fashion relative to the functions for the neutral adaptation condition. The achromatic function changes too, but relatively very little. These theoretically computed adaptation changes are consistent with the kinds of change known to occur in situations of this sort. If the eye that has been adapted to white light is exposed for some time to a saturated red equal in brightness to the white, the normally red end of the spectrum does not become excessively dark, but the amount of redness seen is strongly reduced, and the greens become greatly supersaturated (3, pp. 133-137). Also, the wave length that formerly appeared pure yellow is now strongly greenish, and this is also true for the wave length that formerly appeared pure blue. These changes can be determined from the functions shown in Fig. 12 that have been computed for a given red adaptation, in comparison with the functions for the neutral state that were given in Fig. 6.

From this new set of "adapted" opponent response functions the hue and saturation coefficients and the discrimination data for this new state can also now be determined (26).

These "adapted" response functions are specified, as we said above, for a circumscribed set of conditions for which the photochemical adaptation changes could be taken as primary. As

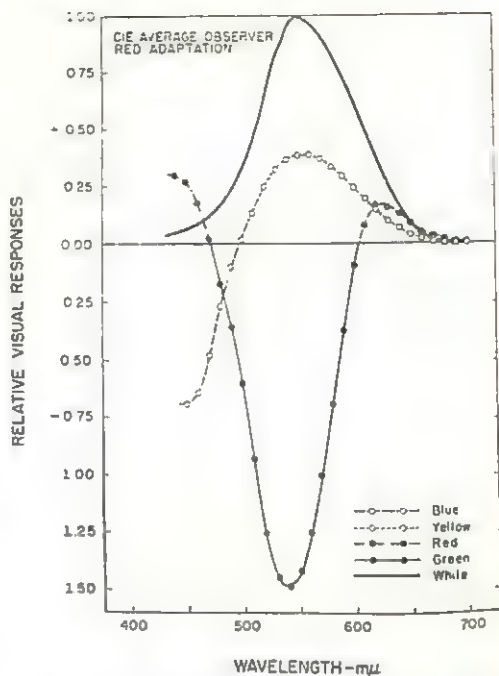


FIG. 12. Chromatic and achromatic visual response functions for red adaptation as predicted by theory.

soon as the relative luminance conditions are altered, however, then the neural inductions enter importantly into the complex picture. For example, if a test stimulus seen within a strongly colored (say, red) surround looks neutral when its luminance is the same as that of the surround, then it will probably appear somewhat reddish at a higher luminance, and the complementary green at a lower luminance (12). (The test stimulus is assumed also to be of predominantly long wave length composition.) In terms of opponent inductions this phenomenon is readily understood. If the red process excited by the red surround induces an opponent green process in the test area, then at an intermediate luminance this green induction is just strong enough to cancel the red-process activity aroused by the test stimulus itself. When the test stimulus is made brighter and the red response to it increases, the unchanged



green induction from the surround becomes inadequate to cancel completely the increased red response to the stronger test stimulus, and the red test hue is now seen. At a much lower luminance of test stimulus, the red process is activated to a much lesser extent, and the green induction from the surround, which is still unchanged in strength, is now sufficient to impart a green hue to the perceived test area. These phenomena are not only consistent with the opponent induction postulate, but they also make it clear why attempts to treat the problem of chromatic adaptation exclusively as a matter of photochemical bleaching are foredoomed to failure (e.g., 1, 33).

### *Color Anomalies and Color Blindness*

When we come to consider individuals who do not have normal color vision we find that their color vision can depart from the normal in two general ways. Their color perceptions may be distorted relative to the normal, or they may exhibit specific color weaknesses or losses. Also, they may show both types of these deviant characteristics at the same time. By distorted color perceptions we mean, for example, the perceptions of the particular type of anomalous individual who has the following characteristics: he sees a distinct orange in the spectral region described normally as pure yellow or nearly so; he needs three stimuli for color mixture; he makes color matches with high precision but uses quite different proportions of the mixture stimuli than does the normal observer. An individual of this type does not seem to have lost any of the efficiency of his neural visual response processes, and it seems reasonable to assume that his color distortions have their basis in the photochemical complex responsible for selective light absorption.

The particular assumptions that we

have made concerning the kinds of deviation that the photosensitive materials may exhibit stem from a generalization made by Dartnall (2), on the basis of his researches concerned with the identification of visual photopigments in a variety of lower organisms. Dartnall has found that when the absorption curves of the various visual pigments are plotted as a function of the vibration frequency of the incident light (the reciprocal of the more usual wave-length specification), all the absorption curves have very nearly the same shape, and they can be made to coincide simply by shifting the curves so that they all reach an absorption maximum at the same frequency. In other words, a single template representing amount of absorption as ordinate, against frequency of radiant energy as abscissa, can be used to fit the absorption function of any visual pigment, whatever the locus of its absorption maximum. It seems reasonable to expect that this same generalization will apply to the photosensitive distributions of anomalous individuals with respect to the population of observers with normal color responses. We have consequently assumed that, in congenital abnormalities of the visual system, the normal photopigments can undergo changes that result in a uniform shift of the entire set of photosensitive distribution functions as a group along the frequency scale. These shifts are assumed to occur in either of two directions: toward higher frequencies (shorter wave lengths) resulting in the type of anomalous color vision identified as *protanomaly*, or toward lower frequencies (longer wave lengths) relative to the normal absorption loci, resulting in the second major type of anomalous color vision known as *deuteranomaly*. The amount of these displacements may also vary in different degrees of congenital anomaly.

Since the absorption of light by the photosensitive materials provides the stimulus for the neural chromatic and achromatic response systems, the visual response functions thus controlled by the deviant photosensitive materials will necessarily be altered, too, and in a systematic manner. Examples of theoretically derived anomalous response functions based on these assumptions are given in Fig. 13. The set of functions in the center block are those for the observer with normal photosensitive materials; those in the upper block are for a protanomalous type whose visual pigment absorptions are assumed to be shifted toward the shorter wave lengths by an amount equal to about 15  $m\mu$

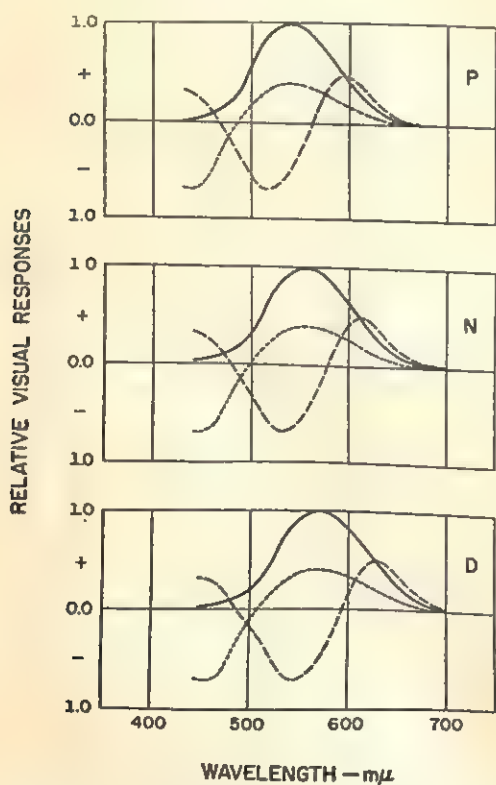


FIG. 13. Theoretical chromatic and achromatic response functions for equal energy spectrum. For observers with protanomalous, normal, and deuteranomalous photoreceptor systems and with normal strength visual response processes.

from the normal peak of about 550  $m\mu$ . This type of individual will have a luminosity function (described by the achromatic, white response function) that peaks at a shorter wave length than the normal and will show considerable loss of luminosity at the red end of the spectrum (48, Ch. 25). The spectral hues will also be altered, with a distinctly reddish yellow occurring where the normal individual sees a unique or pure yellow, whereas the protanomalous observer's pure yellow occurs at a wave length described by the normal as quite greenish. In making color matches, such as a match between 589  $m\mu$  on one side of a bipartite field and a mixture of 530  $m\mu$  and 670  $m\mu$  on the other, this observer will require a much greater proportion of 670  $m\mu$  in the mixture than will the average observer with normal color vision (27, 46). This particular match, the Rayleigh equation, is the earliest and best known diagnostic test for anomalous color vision. In this same test, the anomalous individual whose response functions are shown in the lower block in Fig. 13 will deviate from the normal in the opposite way; that is, he will require a much greater proportion of 530  $m\mu$  in the mixture for the Rayleigh equation (46). This type of anomalous individual (deuteranomalous) is assumed to have photopigment absorptions that are shifted toward the longer wave lengths, and he will see greenish-yellows where the normal sees yellow, yellows where the normal sees orange, etc. Since the neural response processes of both types of anomalies of this sort are assumed to be operating at the normal efficiency, these individuals will show high precision in making their distorted color matches, and their discriminatory capacities will also be good. As a matter of fact, anomalous individuals of this sort have understandably high confi-

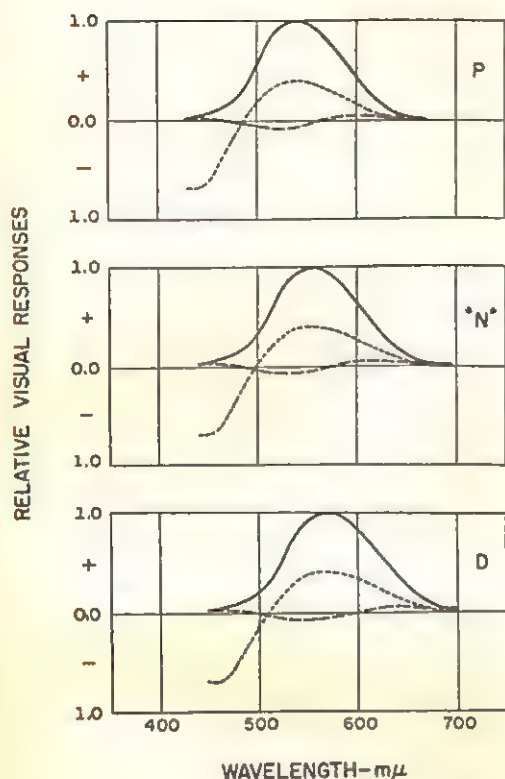


FIG. 14. Theoretical chromatic and achromatic response functions for equal energy spectrum. For observers with protanomaly, normal, and deuteranomaly photoreceptor systems, and with impaired red-green response processes.

dence in their own color capability, and they are extremely resistant toward accepting the results of diagnostic tests which indicate that their visual capacities are deviant from (with the implication of "inferior to") those of the normal population (36, pp. 235-238).

Not all anomalous individuals are as fortunate as the types shown in Fig. 13, however. Many give evidence of real color weakness, in addition to distortions of the kinds already discussed (40). These color-weak individuals seem to have been deprived of some of the efficiency of the neural response processes, particularly of the red-green opponent pair, and their systems may

be represented in terms of the theory by the kinds of response functions given as examples in Fig. 14. The visual pigments of these three types of individuals are taken to be the same as those shown in the preceding figure, respectively, but the red-green paired system is reduced to one-tenth of the normal strength. Such observers have real losses in color discrimination in addition to possible color distortions, and their color matches are imprecise as well as deviant. Individuals with congenitally abnormal color systems are frequently of this general type, and cases of acquired color blindness caused by degenerative disease invariably show this kind of color weakness at some stage in the development of the neural disorder (31).

When the weaknesses become extreme, whether in congenital or acquired disorders, the red-green system seems to be entirely lost to normal function, and a condition of dichromasy, or so-called "color-blindness," results. That is, the visual system becomes a two-variable

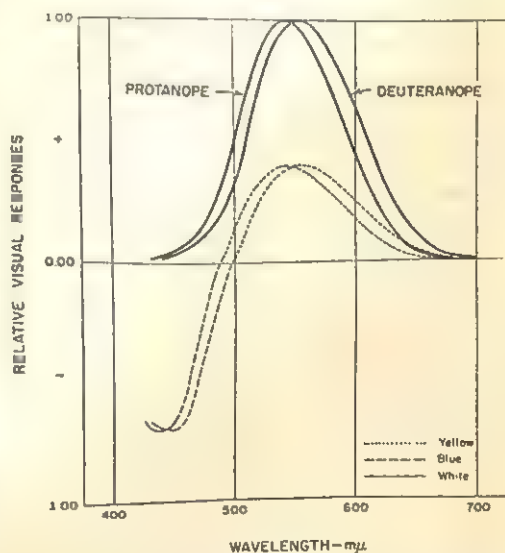


FIG. 15. Theoretical chromatic and achromatic response functions for equal energy spectrum. For observers with nonfunctioning red-green response processes.



one, as shown in Fig. 15. Here the yellow-blue and the white-black neural systems remain intact and functioning, but there is no red-green response function. If the red-green loss occurs without changes in the visual pigments, the remaining yellow-blue and white-black response functions are like those of the normal individual; but, since there is no red-green system, the spectrum is divided into only two hue sections for these individuals. The short wave lengths which normally vary from violet through blue and blue-green to pure green all appear as blue, but of varying saturations, with a neutral region where the normal pure green occurs. Beyond this wave length the remainder of the spectrum appears yellow, in varying saturations, out to the extreme long-wave limit of visibility. The luminosity function is the same as for the observer with normal color vision. Individuals who fit this response pattern would be classified as *deuteranopes* (29). If the visual pigments are altered, so as to produce an absorption shift toward the short wave lengths in addition to the complete red-green neural loss, then the spectrum is again divided into a short-wave blue and a long-wave yellow section, but the neutral region that divides the spectrum into the two major hues occurs at a shorter wave length than for the deuteranopes. The luminosity function is also displaced in this type of dichromasy, as it is for the anomalous individuals with similar photopigment changes, and the type of "color-blind" vision associated with this pattern is called *protanopia* (29).

These two theoretically assumed kinds of deviation from the normal system—i.e., photopigment changes and neural losses or weaknesses of the paired red-green response system—permit us to assemble a systematic picture of the many various manifestations of abnor-

mal red-green vision that defy understanding in terms of any model of the visual system that assumes a one-to-one correspondence between light absorption in the retinal receptors and the resulting color sensations (22, 27).

Defects or losses may also occur in the yellow-blue neural response system, although such defects seem to be much more rare than the red-green defects. Again, these yellow-blue neural losses may take place either with or without changes in the photosensitive materials in the retina. Examples of the remaining red-green and white-black response functions in two types of yellow-blue blindness are given in Fig. 16. In each type of this disorder, the yellow-blue neural response function is missing, and the total gamut of colors for these individuals includes only neutral and reds and greens of various saturations. If there is no simultaneous photopigment disorder, there are two neutral points in the spectrum, one in the region where the normal sees a pure yellow, and an-

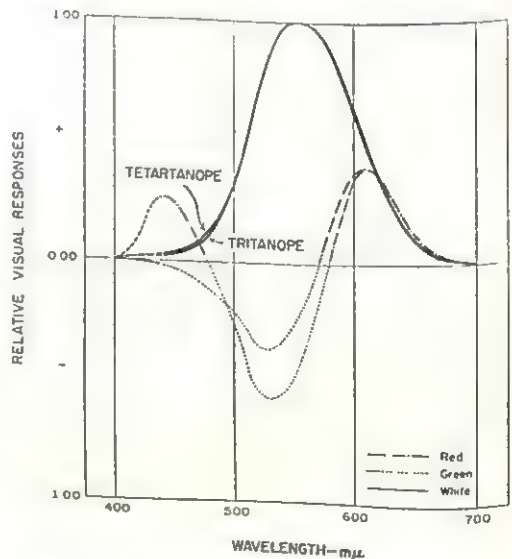


FIG. 16. Theoretical chromatic and achromatic response functions for equal energy spectrum. For observers with nonfunctioning yellow-blue response processes.

other in the region where the normal sees a pure blue. Yellow-blue blindness of this sort is called *tetartanopia*, and only a few cases of it have been reported in the literature (e.g., 34, pp. 68-92). Slightly more common is the second type of yellow-blue blindness, known as *tritanopia* (49), in which not only the neural yellow-blue system is lost, but also the short-wave photopigment seems to be missing. Observers of this type have a neutral point in the normally yellow-green region of the spectrum, but there is no second neutral point, and the green hues extend into the short-wave region that appears violet to the person with normal color vision.

For all these types of deviant color vision, calculation from the theoretical spectral response functions of discrimination curves, color mixture equations, and other psychophysical relations are in good agreement with the experimental data that are available for the various kinds of defective color systems (22, 27).

### *Opponents-Theory and Neurophysiology*

The conceptual model for the opponent-colors theory as originally presented by Hering drew its sharpest criticism on the grounds of being bad physiology. Some of this criticism was based on an erroneous interpretation of Hering's views, an interpretation that incorrectly assigned the opponent processes to the photochemical activities in the retinal cells. Hering's own concept of mutually opponent neural processes, each capable of being activated by external stimulation, was also, however, far ahead of the knowledge of neurophysiology at the time it was proposed (16). But this concept now turns out to be perfectly consistent with the picture of neural function that is only just recently beginning to build up from elec-

trophysiological studies of the visual neural apparatus.

It has become clear that nerves do not simply respond or fail to respond when a stimulus is presented to the appropriate end-organ. Rather, they may respond according to any of a number of quite specific patterns. For example, a nerve fiber may (a) discharge at the onset of stimulation and subsequently gradually become quiet; (b) discharge at both onset and cessation of stimulation with a quiet period in between; or (c) cease any spontaneous activity when first stimulated and during continued stimulation, but respond with a burst of electrical impulses when the stimulus ceases to act (7). The on- and off-phases of discharge are mutually inhibitory processes, they are associated with slow electrical potentials of opposite sign, and they cancel each other when the experimental conditions are so manipulated as to cause both on- and off-discharges to impinge simultaneously on the same ganglion cell (6). In Granit's opinion (6), the evidence from electrophysiology provides a "belated vindication of Hering's view" that the visual system is characterized by mutually opponent neural processes.

The concept of mutual interaction among the various elements of the physiological field is also basic to the theory and is critical to an understanding of both areal effects and simultaneous contrast phenomena. Here again, we find the researches in electrophysiology indicating that individual nerve elements never act independently, and that visual function must be thought of in terms of the integrated action of all the units of the neural visual system (8). Hartline (9) has found that, even in the very simple *Limulus* eye, the discharge of impulses in any one optic nerve fiber depends not only upon the stimulus to the specific receptor unit



from which that fiber arises but also upon the stimulation over the entire population of mutually interacting elements. Both excitatory and inhibitory interactions of the sort to be expected by theory have actually been demonstrated in the neural responses of the vertebrate visual system by Hartline (8), Kuffler (32), and Granit (6).

The way in which the postulated three independent systems of paired opponent processes ( $y-b$ ,  $r-g$ ,  $w-bk$ ) are differentiated neurally is still a matter for conjecture. Hering thought it was a matter of process specificity, but was willing to use the concept of material, or structural, specificity, which he guessed would be more readily comprehended by most interested readers of his views at the time. Our own theoretical preference at this time is the conjecture that a particular color quality is more probably determined by a particular state of the nervous tissue than by activity of a particular structural element in the nervous network. Thus, we would be inclined to look for a difference between yellow-blue vs. red-green processes, rather than toward isolation of yellow-blue or red-green fibers or nerve cells.

#### SUMMARY

This paper has presented a summary of our progress to date in providing a quantitative formulation for the Hering opponent-colors theory, and in relating the postulated visual mechanism to specific problems of color sensation, color mixture and color discrimination; to the dependence of these functions on the physical variables of both stimulus wave length and energy level; to their further dependence on adapting and surround stimulation; and to the changes in these functions that occur in various kinds of abnormal color vision. It is

our conclusion that the opponent-colors theory serves as a fruitful working hypothesis by bringing a systematic coherence to the mass of isolated color phenomena that have been reported and subjected to quantitative experiment throughout the years. The physiological concepts basic to the theory are also shown to be consistent with recent findings in neurophysiology.

#### REFERENCES

1. BREWER, W. L. Fundamental response functions and binocular color matching. *J. Opt. Soc. Amer.*, 1954, 44, 207-212.
2. DARTNALL, H. J. A. The interpretation of spectral sensitivity curves. *Brit. med. Bull.*, 1953, 9, 24-30.
3. EVANS, R. M. *An introduction to color*. New York: Wiley, 1948.
4. FARNSWORTH, D. Tritanomalous vision as a threshold function. *Die Farbe*, 1955, 4, 185-196.
5. GRAHAM, C. H. Vision: III. Some neural correlations. In C. Murchison (Ed.), *A handbook of general experimental psychology*. Worcester: Clark Univer. Press, 1934. Pp. 829-879.
6. GRANIT, R. *Receptors and sensory perception*. New Haven: Yale Univer. Press, 1955.
7. HARTLINE, H. K. The response of single optic nerve fibers of the vertebrate eye to illumination of the retina. *Amer. J. Physiol.*, 1938, 121, 400-415.
8. HARTLINE, H. K. The neural mechanisms of vision. *Harvey Lectures*, 1941-42, 37, 39-68.
9. HARTLINE, H. K., WAGNER, H. G., & RATLIFF, F. Inhibition in the eye of limulus. *J. gen. Physiol.*, 1956, 39, 651-673.
10. HARTRIDGE, H. The polychromatic theory. *Documenta Ophthal.*, 1949, 3, 166-193.
11. HECHT, S. Vision: II. The nature of the photoreceptor process. In C. Murchison (Ed.), *A handbook of general experimental psychology*. Worcester: Clark Univer. Press, 1934. Pp. 704-828.
12. HELSON, H. Fundamental problems in color vision. I. The principle governing changes in hue, saturation, and lightness of non-selective samples in



- chromatic illumination. *J. exp. Psychol.*, 1938, 23, 439-476.
13. HERING, E. *Zur Lehre vom Lichtsinne*. Berlin, 1878.
14. HERING, E. *Zur Erklärung der Farbenblindheit aus der Theorie der Gegenfarben*. *Lotus, Jb. f. Naturwiss.*, 1880, 1, 76-107.
15. HERING, E. *Ueber Newton's Gesetz der Farbenmischung*. *Lotus, Jb. f. Naturwiss.*, 1887, 7, 177-268.
16. HERING, E. *Zur Theorie der Vorgänge in der lebendigen Substanz*. Prague: 1888. (English translation by F. Welby, in *Brain*, 1897, 20, 232-258.)
17. HERING, E. *Grundzüge der Lehre vom Lichtsinn*. Berlin: Springer, 1920.
18. HOCHBERG, J. E., TRIEBEL, W., & SEAMAN, G. Color adaptation under conditions of homogeneous visual stimulation (Ganzfeld). *J. exp. Psychol.*, 1951, 41, 153-159.
19. HURVICH, L. M., & JAMESON, DOROTHEA. The binocular fusion of yellow in relation to color theories. *Science*, 1951, 114, 199-202.
20. HURVICH, L. M., & JAMESON, DOROTHEA. Spectral sensitivity of the fovea. I. Neutral adaptation. *J. Opt. Soc. Amer.*, 1953, 43, 485-494.
21. HURVICH, L. M., & JAMESON, DOROTHEA. A quantitative theoretical account of color vision. *Trans. N. Y. Acad. Sci.*, 1955, 18, 33-38.
22. HURVICH, L. M., & JAMESON, DOROTHEA. Some quantitative aspects of an opponent-colors theory. II. Brightness, saturation, and hue in normal and dichromatic vision. *J. Opt. Soc. Amer.*, 1955, 45, 602-616.
23. HURVICH, L. M., & JAMESON, DOROTHEA. Some quantitative aspects of an opponent-colors theory. IV. A psychological color specification system. *J. Opt. Soc. Amer.*, 1956, 46, 416-421.
24. ISHAK, I. G. H. Determination of the tristimulus values of the spectrum for eight Egyptian observers and one British observer. *J. Opt. Soc. Amer.*, 1952, 42, 844-849.
25. JAMESON, DOROTHEA, & HURVICH, L. M. Some quantitative aspects of an opponent-colors theory. I. Chromatic responses and spectral saturation. *J. Opt. Soc. Amer.*, 1955, 45, 546-552.
26. JAMESON, DOROTHEA, & HURVICH, L. M. Some quantitative aspects of an opponent-colors theory. III. Changes in brightness, saturation, and hue with chromatic adaptation. *J. Opt. Soc. Amer.*, 1956, 46, 405-415.
27. JAMESON, DOROTHEA, & HURVICH, L. M. Theoretical analysis of anomalous color vision. *J. Opt. Soc. Amer.*, 1956, 46, 1075-1089.
28. JONES, L. A., & LOWRY, E. M. Retinal sensibility to saturation differences. *J. Opt. Soc. Amer.*, 1926, 13, 25-34.
29. JUDD, D. B. Current views on colour blindness. *Documenta Ophthalmol.*, 1949, 3, 251-288.
30. JUDD, D. B. Basic correlates of the visual stimulus. In S. S. Stevens (Ed.), *Handbook of experimental psychology*. New York: Wiley, 1951. Pp. 811-867.
31. KÖLLNER, H. *Die Störungen des Farbensinnes*. Berlin: S. Karger, 1912.
32. KUFFLER, S. W. Discharge patterns and functional organization of mammalian retina. *J. Neurophysiol.*, 1953, 16, 37-68.
33. MACADAM, D. L. Chromatic adaptation. *J. Opt. Soc. Amer.*, 1956, 46, 500-513.
34. MÜLLER, G. E. *Darstellung und Erklärung der verschiedenen Typen der Farbenblindheit*. Göttingen: Vandenhoeck and Ruprecht, 1924.
35. PARSONS, J. H. *An introduction to the study of colour vision*. (2nd ed.) Cambridge: Cambridge Univer. Press, 1924.
36. PICKFORD, R. W. *Individual differences in colour vision*. London: Routledge and Kegan Paul, 1951.
37. PIÉRON, H. La dissociation de l'adaptation lumineuse et de l'adaptation chromatique et ses conséquences théoriques. *Année psychol.*, 1939, 40, 1-14.
38. PURDY, D. M. The Bezold-Brücke phenomenon and contours for constant hue. *Amer. J. Psychol.*, 1937, 49, 313-315.
39. RIGGS, L. A., RATLIFF, F., CORNSWEET, JANET C., & CORNSWEET, T. N. The disappearance of steadily fixated visual test objects. *J. Opt. Soc. Amer.*, 1953, 43, 495-501.
40. ROSMANIT, J. *Anleitung zur Feststellung der Farbensichtigkeit*. Leipzig: Deuticke, 1914.
41. TROLAND, L. T. Apparent brightness: its conditions and properties. *Trans. Illum. Engr. Soc.*, 1916, 11, 957-966.
42. TROLAND, L. T. *The principles of psychophysiology*. Vol. 2. *Sensation*. New York: D. Van Nostrand, 1930.

43. VON BEZOLD, W. Ueber das Gesetz der Farbenmischung und die physiologischen Grundfarben. *Ann. Phys. u. Chem.*, 1873, 150, 221-247.
44. VON KRIES, J. Die Gesichtsempfindungen. In W. Nagel (Ed.), *Handbuch der Physiologie des Menschen*. Brunswick: Vieweg, 1905. Pp. 109-282.
45. WEALE, R. A. Hue-discrimination in para-central parts of the human retina measured at different luminance levels. *J. Physiol.*, 1951, 113, 115-122.
46. WILLIS, MARION P., & FARNSWORTH, D. Comparative evaluation of anomalous scopes. *Med. Res. Lab. Rep.* No. 190, 1952, 11, No. 7, 1-89.
47. WRIGHT, W. D., & PITT, F. H. G. The colour-vision characteristics of two trichromats. *Proc. Phys. Soc. (London)*, 1935, 47, 205-217.
48. WRIGHT, W. D. *Researches on normal and defective colour vision*. St. Louis: Mosby, 1947.
49. WRIGHT, W. D. The characteristics of tritanopia. *J. Opt. Soc. Amer.*, 1952, 42, 509-521.

(Received for early publication July 1, 1957)

















